*Research Article*

# Analysis of Pedestrians' Road Crossing Behavior, in Social Groups

**Leidy Barón[1]** (iD)**, Susana Faria[2], Emanuel Sousa[3]** (iD)**,
and Elisabete Freitas[1]**

## Abstract

A significant percentage of pedestrians walk in social groups (friends, families, or acquaintances who walk together). Although patterns generated by social interactions among group members have been shown to affect crowd dynamics, studies on the effect of social interactions at different crossing phases under low pedestrian density are limited. This study aims to comprehensively examine the influence of size and sex composition on pedestrians' behaviors when walking alone and with friends in different phases before, during, and after the road crossing. For this, experiences were carried out with controlled small groups of friends (varying size and sex composition) at three unsignalized crosswalks with low pedestrian density. The average speed and distance between the young pedestrians in six segments of the trajectories (two in each phase), extracted from video recordings, were analyzed with linear mixed models. Results show that pedestrians reduce their speed when approaching the curb, they accelerate while on the crosswalk, and reduce again when they reach the other side. In all phases, the average speed of the groups was lower than the single pedestrians, and the females' groups walked slower than the males, except during the crossing, where no sex-related differences were found. On the contrary, before the crossing, the distance increased and decreased from the second segment in the crosswalk. The smallest distance was observed between the female groups and dyads. These findings have relevant implications for research on pedestrian behavior, helping to better understand the complexity of pedestrian dynamics and improve pedestrian safety.

## Keywords

pedestrian behavior, social groups, crossing behavior, pedestrian safety

Social interaction plays a significant role in shaping the walking behavior of pedestrians, particularly on crosswalks or on sidewalks where they frequently have to adjust their path to accommodate others. Furthermore, pedestrians often walk in small social groups, comprising individuals with some kind of social connection, such as couples, friends, or families (*1–4*). The coherent movement of these groups is influenced by social rules that impact their walking dynamics. They share the same destiny and intentionally walk together, avoiding splitting (*5, 6*). Social groups within a pedestrian flow can range from two to six individuals, but most frequently have only two or three. Groups with more than four members have a tendency to split into single individuals, dyads, and triads (*7–9*). A possible reason for this is that in big groups, people tend, for the sake of efficiency, to focus communication only on the surrounding elements (*3*). It

has been reported that the formations they adopt and their walking speeds facilitate communication, even though this may hinder the task of moving through a flow of pedestrians (*3, 10, 11*).

The behavior of social groups has been mainly analyzed in two scenarios: (i) in normal conditions (i.e., non-emergency and non-mass events), such as crowded commercial walkways or corridors connecting a train station with a shopping center (*3, 6, 9, 12*), and (ii) more complex

[1]Institute for Sustainability and Innovation in Structural Engineering (ISISE), University of Minho, Guimarães, Portugal
[2]Centre of Mathematics (CMAT), University of Minho, Guimarães, Portugal
[3]Department of Human-Technology Interaction and Robotics (HTIR), Center for Computer Graphics, Guimarães, Portugal

**Corresponding Author:**
Leidy Barón, leidymbaron95@gmail.com

scenarios, such as during evacuation processes or mass events (*13–16*). In most of these studies, the observed difference between individuals and groups was related to the need of the latter ones to maintain cohesion, to facilitate social interactions and communication (*3, 17*). Notably, the majority of them were conducted under medium to high population densities, and thus tthe results obtained could have been explained mostly by the effort required for the groups to stay together. Most importantly, to our knowledge, no study has ever addressed crossing behavior. This is an important gap in pedestrian behavior research, considering the prevalence of group walking and noting that road crossing remains one of the most safety-critical moments in everyday urban walking.

Pedestrian behavior during interaction with vehicles is critical because it directly affects the likelihood and severity of accidents. The high number of pedestrian crashes at unsignalized intersections and the likelihood of severe injury or death owing to pedestrian vulnerability have led to extensive research aimed at understanding how pedestrian–vehicle interactions occur at such crosswalk locations (*18*). For example, different behavioral factors that could lead to pedestrian–vehicle collisions have been investigated, including drivers' compliance with pedestrians' right of way on zebra crossing and age-driven pedestrian crossing behavior, focusing on motor skills (speeds and trajectories) and decision-making (accepted time gap to cross) (*19*). Factors affecting pedestrians' crosswalk decisions have been investigated such as position and speed of the approaching vehicle (*20, 21*) or the influence of the built environment (perception of comfort and safety) (*22*). Quite strikingly, the behavior of social groups when approaching road crossings remains largely unexplored. The goal of this research is thus to explore walking and road crossing behavior of social groups of different sizes and sex compositions. The study is focused on urban contexts in streets with low pedestrian density and in normal conditions.

Pedestrian group walking behavior can be categorized according to walking speed and spatial organization. To interpret these variables correctly, three interrelated levels of analysis must be considered: the individual level (personal factors), the external level (environmental factors), and the group level (social factors) (*1, 23, 24*).

Concerning personal factors, some studies have related demographic characteristics, such as age and sex, to variations in walking speeds. For example, several studies have shown that walking speed declines, on average, with increasing age (*25–28*). Moreover, on average, men walked faster than women (*8, 27*).

Beyond personal dimensions, external factors associated with the environment where pedestrians walk, can also influence walking behavior. These include the geographic locations, cultural and traffic regulations, infrastructure characteristics, the presence of obstacles, and traffic conditions (*23, 27, 29*). As regards social factors, there is substantial evidence that variables such as group size and sex composition influence the walking speed of groups of pedestrians. Social groups are on average slower than isolated individuals and speed decreases as the group size increases (*27, 30, 31*). This may happen because individuals in social groups become distracted while talking to each other, or they may adjust their speed to the slowest member (*3*). Nessler and Gilliland (*32*) suggest that people with similar leg lengths, when walking side by side tend to unconsciously synchronize their stepping, which can, in turn, affect their walking speed. There is also evidence that the speed of pedestrians' groups significantly depends on the surrounding pedestrian density level, with people in groups walking faster at lower densities (*6*). Concerning sex, there seems to be a tendency for male groups to walk more "scattered" than female or mixed groups, and walking speed was highest for male triads compared with mixed and all-female triads (*7, 27*).

Researching pedestrian crossing behavior can be challenging owing to the multitude of factors that influence it and that are difficult to control. Many studies found in the literature use field observations as a data collection method. These provide the most naturalistic results about the movement and behavior of street users. However, crossing situations in realistic settings are highly variable. Many observations are necessary to obtain statistically significant conclusions, making data collection and analysis time-consuming. Additionally, individual characteristics of pedestrians are difficult to capture with this method (*23*). Contrarily, the effect of particular factors can be easily observed and analyzed using controlled experiments. These are predominantly used to study the operational movement behavior of pedestrians. However, it is often challenging to ensure that data collected in a laboratory can represent real-life pedestrian behavior or be transferable to different situations (*23*). Semi-controlled studies may be a valid compromise between controlled and naturalistic approaches. These experiments take place in real settings, but the researchers can assign participants to groups, select the route, and define the tasks, while participants have freedom to decide if, when, where, and how to perform them (*33*).

The complexity of pedestrian behavior also poses challenges for data analysis. To address potential unobserved heterogeneity across different study sites and individual pedestrians, several pedestrian behavior studies have employed multilevel or hierarchical statistical models in field observations or controlled experiments with repeated observations from each participant. These models include frequentist or Bayesian frameworks such as
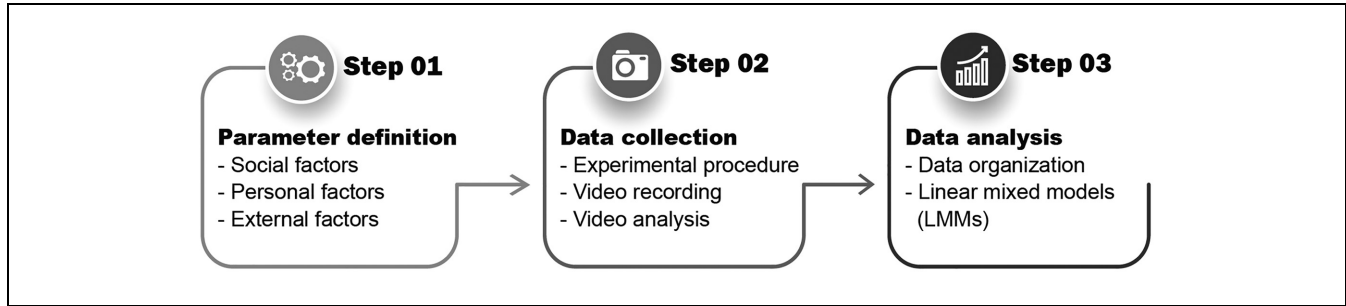
**Figure 1.** Methodology process steps.

linear mixed models (LMM), generalized linear mixed models (GLMM), and Bayesian multilevel models (BMM) (*34*, *35*). The choice of statistical framework depends on various factors, such as the research question, data, model complexity, and researcher familiarity with each framework (*36*). Frequentist frameworks are suited for scenarios where computational efficiency is relevant, simpler models are appropriate, well-established normal distributions exist, or when researchers are more familiar with these methods. Bayesian frameworks are preferable when modeling complex relationships and temporal dependencies, incorporating prior knowledge, non-normal distributions, or complex random effects structures (*35–37*). In this study, we selected a frequentist approach, specifically LMMs, as they are aligned with the data characteristics and research question and have proven effective in modeling multilevel data on pedestrians' crossing behaviors (*34*).

In this sense, this work intends to study the influence of size and sex composition on pedestrian behavior when walking in a group in urban areas in normal conditions with low pedestrian density and the difference in the pedestrian behaviors of the groups in three phases: before, during and after crossing the road. For this, semi-controlled experiences were carried out in three urban sites in Guimarães, with small groups of young pedestrians (1, 2, and 3) who had a defined path to walk, varying group size and sex composition.

The work contributes to an improved understanding of how speed and distance between pedestrians vary at different phases of street crossings, when walking in social groups. The resulting insights are relevant to improve simulation techniques, optimize pedestrian facilities, and assign more realistic values in modeling the operational behavior of pedestrians to develop effective strategies for reducing accidents and improving safety for road users.

## Method

The experimental approach used in this study is divided into three steps: (1) parameter definition; (2) data collection; and (3) data analysis (see Figure 1). These steps are described below.

### Parameter Definition

To define the variables analyzed in this study, the selection of social, personal, and external factors is presented below.

*Social Factors.* Previous studies have identified that group sizes ranging from two to four people are the most common. However, they can be affected by the pedestrian density and population of the study location. To determine the most common group sizes in Guimarães, preliminary counts of pedestrian groups were conducted at two sites—Londres and São Gonçalo Avenues. Videos were recorded for a duration of one hour at each location. These streets are characterized by mixed land use and are located close to the city center.

The number of pedestrians per group was counted (a) on the sidewalks and (b) on the crosswalks. The results are presented in Figure 2. As one can see, at least 40% of pedestrians walk in groups, with the most common sizes being 1, 2, and 3 people, while groups of more than three people only represent 4% of the total. Therefore, this study focused on analyzing the behavior of groups of pedestrians consisting of 1, 2, and 3 people (group size). Female, male, and mixed groups were considered based on the participants' personal characteristics. The measurement of the participants' leg length was also considered, as there are indications that it may influence behavioral coordination.

*Personal Factors.* In this study, participants were recruited from the University of Minho community in Portugal. The study consisted of 90 adults, including 45 females and 45 males, with ages ranging from 23 to 43 years old (mean = 27.8 years old; standard deviation = 3.9 years old). All participants provided informed consent before participating in the experiments. Demographic information, such as sex and age, was recorded for each
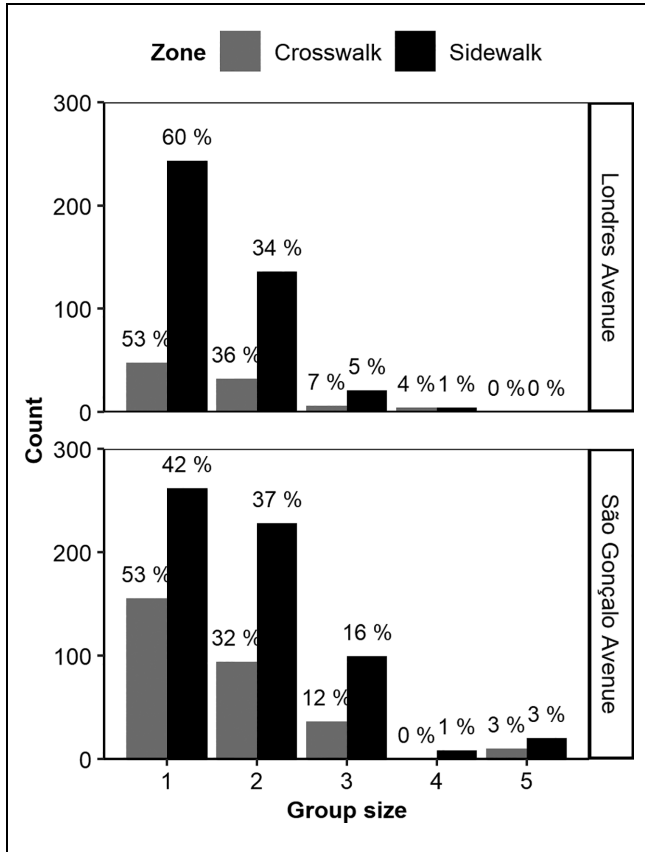
**Figure 2.** Pedestrians group walk on the crosswalks and sidewalks of *Londres* and *São Gonçalo* avenue.

participant, and leg length was also measured before the experiments began (minimum = 0.82 m, maximum = 1.20 m, mean = 0.95 m, standard deviation = 0.078 m). Since all participants were young adults within the same age range, this factor was not analyzed in the study owing to the lack of variability.

*External Factors.* The experiments were conducted in unsignalized crosswalks in an urban area, namely on

three streets in Guimarães, Portugal: *Teixeira de Pascoais* street, *Martins Sarmento* square, and *Doutor Joaquim de Meira* street (see Figure 3).

The three streets being studied have varying sidewalk and crosswalk widths, as well as different numbers and placements of obstacles. Each street is a collector road without traffic lights, with two lanes, and mixed land use that includes both commercial and residential properties. A reasonable volume of pedestrian traffic is observed on these streets, resulting in a level of service (LOS) A on the sidewalks according to the Highway Capacity Manual (HCM) 2010 reference values (*38*). Table 1 presents the main characteristics and features of the streets. The infrastructure characteristics were measured on-site.

To characterize the space occupied by obstacles on the sidewalks, the effective width was calculated according to HCM, that defines it as the total walkway width minus the width of unusable buffer zones along the curb and building line (HCM, 2010). The widths and shy distances from obstructions considered are presented in Appendix A.

### Data Collection

*Experimental Procedure.* For the experiment, participants were asked to walk a fixed trajectory, from predefined point A to B and vice versa (returning from point B to point A). The circuits included an initial part where participants walk on the sidewalk, followed by a road cross in a crosswalk, then back to the sidewalk on the other side of the road (see Figure 4). At the end of the trajectory, the pedestrians were instructed to return to the starting point, executing the inverse path.

For each experiment, a cluster of five participants was requested to participate. Half of the clusters consisted of three males and two females (type 1), while the other part consisted of three females and two males (type 2) (Figure 5a). Each cluster was then organized into six experimental groups formed by either a single person, two dyads, or three triads. To analyze the influence of



**Figure 3.** Study section: (*a*) site 1—*Teixeira de Pascoais* Street, (*b*) site 2—*Martins Sarmento* Square, (*c*) site 3—*Doutor Joaquim de Meira* Street, Guimarães, Portugal.

**Table 1.** Main Characteristics of the Study Sites

| Characteristics | Study site | | |
| --- | --- | --- | --- |
| | Site 1 | Site 2 | Site 3 |
| Road pavement | Asphalt concrete | Cobblestones | Asphalt concrete |
| Distance to a bus stop (m) | 99.13 | 9.75 | 1.2 |
| Park on street | Yes | No | No |
| Width of the crosswalk (m) | 5.12 | 2.983 | 3.994 |
| Length of the crosswalk (m) | 7.13 | 7.994 | 8.790 |
| Average width of the lanes (m) | 2.950 | 3.265 | 2.570 |
| Effective width of the sidewalk A (m) | 8.137 | 2.141 | 1.535 |
| Effective width of the sidewalk B (m) | 9.536 | 1.987 | 1.457 |
| Pedestrian volume (15-min) | 65 | 42 | 28 |
| Pedestrian unit flow rate (p/min/m) | 0.68 | 1.543 | 1.144 |
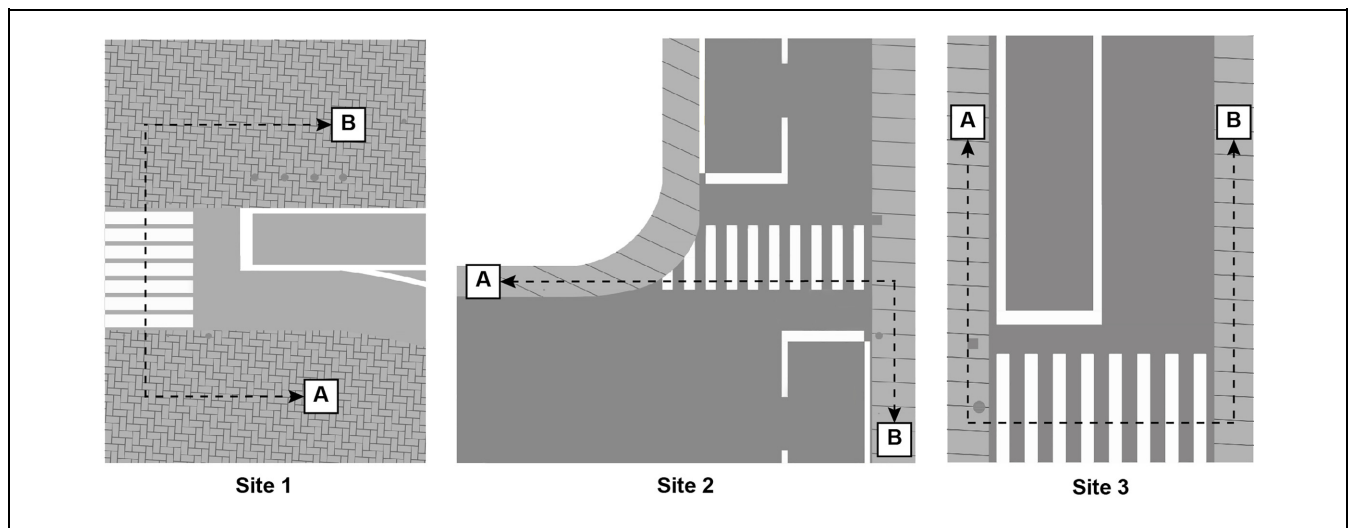


**Figure 4.** Pedestrians' trajectories in study sections: (*a*) site 1, (*b*) site 2, (*c*) site 3.

social factors on individual behavior, one participant was kept fixed as the cluster's reference person and was part of all its groups. The composition of the remaining participants in the group, as regards their sex and number, was varied across the six experimental groups. This made it possible to study how the behavior of the reference person was influenced by different social factors in the group. In cluster type 1, the reference person was a man, while in cluster type 2, it was a woman.

Following the experimental procedure, each of the six small groups walked the predetermined trajectory repeating it twice in each experiment. The experiments lasted approximately 40 min. A total of eighteen clusters took part in the study (six clusters per study site), divided into 18 singles, 24 dyads, and 36 triads.

*Video Recordings.* The experimental trials were recorded using a GoPro Hero5 black video camera with a resolution of 1920 × 1080 pixels, a frame rate of 60 fps, and linear range. The camera was placed at a height between 2.5 and 4 m and between 2 and 10 m away from the crosswalk, depending on the conditions of each observation site.

Following the experiments, he video recordings were analyzed using T-Analyst (Figure 6*a*) (Johnsson, Norén and Laureshyn, 2020). This is a semi-automated video analysis software developed at Lund University and available as freeware (*39*). This tool has been used in several road safety studies (*40–42*).

The software calculates the trajectories (*x* and *y* coordinates) and the speed of the pedestrians. Calibrations were made in a standalone software named T-Calibration using the TSAI-calibration algorithm (*43*) (Figure 6*b*). Manual tracking was chosen to overcome the obstacles that represented potential pedestrian occlusions and reduce the error of distances between the pedestrians walking in groups.

The center of mass coordinates obtained for each participant allowed the calculation of the distance between each pair of pedestrians in the order shown in Figure 7.
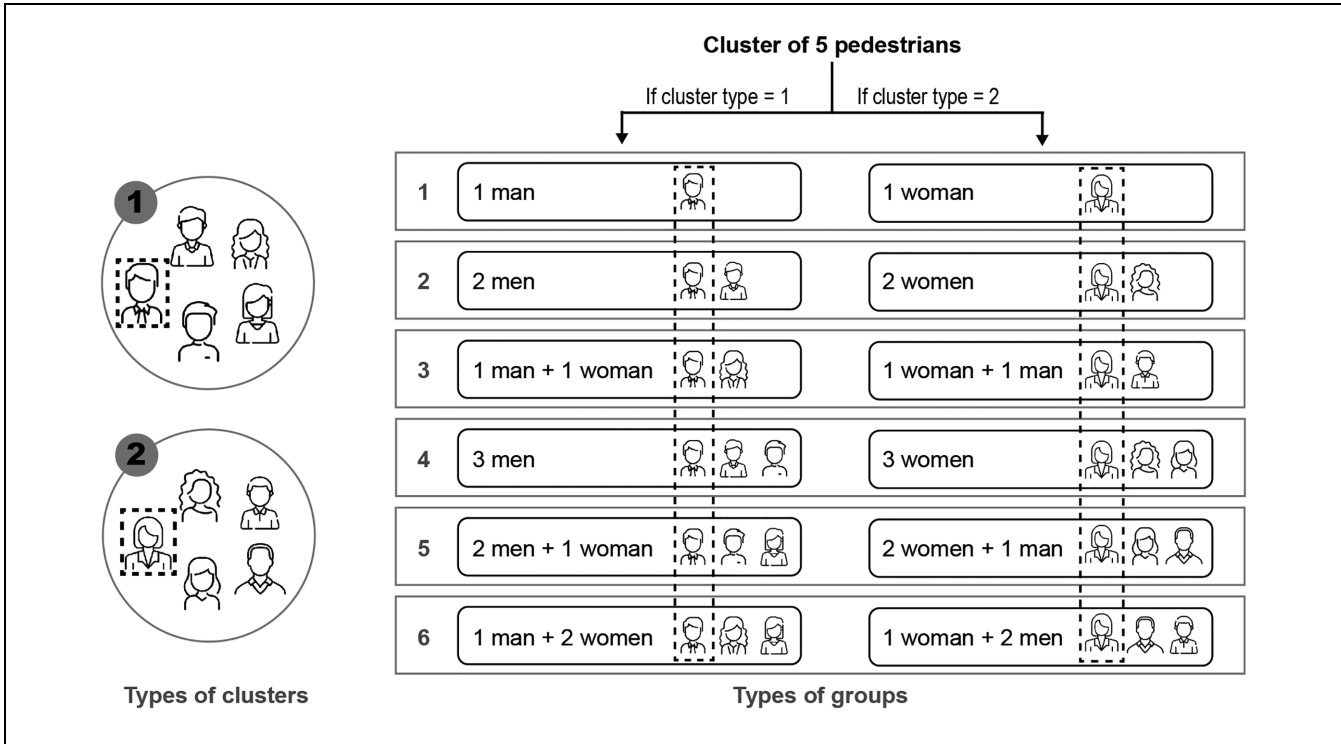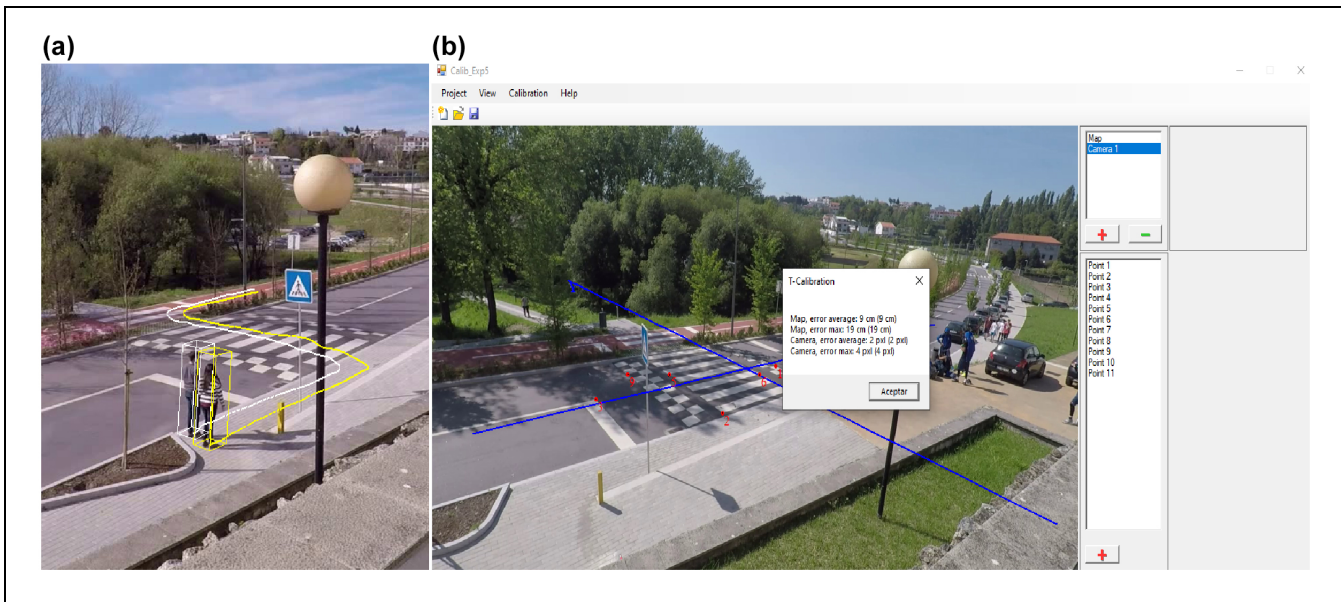
**Figure 5.** Types of clusters and types of groups.



**Figure 6.** Example of pilot experiment: (*a*) tracking using T analyst, and (*b*) calibration using T calibration

## Data Analysis

*Data Organization.* The pedestrians' trajectories were split into three analysis phases: (i) before the crossing, (ii) during the crossing, and (iii) after crossing. The path of analysis is shown in Figure 8, which includes the width of each zone of sidewalks or crosswalks according to the study site. Additionally, it includes the length of 9 m for sidewalks and the length of each crosswalk. Each phase was divided into two segments to study possible differences in the areas closest to the road crossings, before
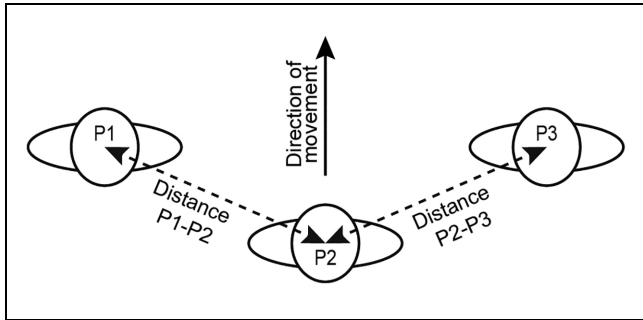
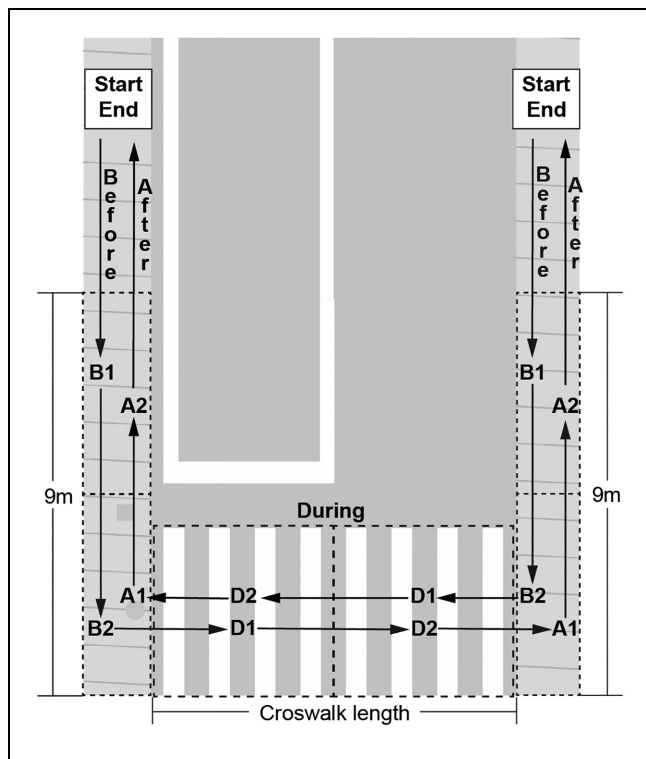**Figure 7.** Distances between pedestrians.



**Figure 8.** Area of analysis: crossing phases divided into segments.

(segment B1, B2), during (segment D1 e D2) and after (segment A1 e A2) the road crossing.

The final dataset includes the dependent and independent variables. The dependent variables, which were calculated in the segments of each phase, were the average walking speed and average distance between pairs of pedestrians. The independent variables correspond to the factors identified in the literature that influence the pedestrians' behavior and were divided according to the social and external factors (see Table 2). Personal characteristics are implicit within social factors (sex of pedestrians and leg length).

Figure 9 shows a summary scheme of the groups of factors considered to study their influence on pedestrian behavior, the variables analyzed, and the analysis phases.

*Modelling Pedestrian Walking Behavior.* Data was modelled using LMMs. This class of models is primarily used to describe linear relationships between a response variable and some covariates in grouped data according to one or more factors. LMMs incorporate fixed effects, which are parameters associated with an entire population or with certain repeatable levels of factors, and random effects, which are associated with individual experimental units drawn at random from a population (44).

LMMs are an appropriate technique for analyzing nested structured data, such as the data presented in this study, owing to repeated measurements. The participants were divided into small groups, and each small group walked twice the predefined trajectory (see section 2.2—Experimental procedure).

LMMs were estimated to assess the influence of all the variables mentioned above (see Table 2) on the distance between pedestrians and the pedestrians' walking speed.

The hierarchical data structures have three levels (see multilevel diagrams Figure 10). Each observation corresponds to a trajectory of a pair of pedestrians (Pair_ID) for the distance between pedestrians and, of each pedestrian (Ped_ID), for the walking speed. These form part of an upper-level group of participants (Cluste_ID).

The model, for $i = 1, \ldots, m; j = 1, \ldots, m\_i, k = 1, \ldots, m\_ij$ can be written as

$$Y_{ijk} = \beta_O + \sum_{p=1}^{P} \beta_p X_{pijk} + \sum_{l=1}^{L} \alpha_l Z_{li} + u_{ij} + b_i + \varepsilon_{ijk}$$
$$u_{ij} \sim N\left(0, \sigma_U^2\right)$$
$$b_i \sim N\left(0, \sigma_I^2\right)$$
$$\varepsilon_{ijk} \sim N\left(0, \sigma_E^2\right)$$

$$(1)$$

where $Y_{ijk}$ is the response variable (pedestrians speed or distance between pedestrians) for the $k$th trajectory from the $j$th pedestrian or pair of pedestrians of the $i$th group of participants, $X_{Pijk}$ is the corresponding value of $p$th predictor variable at trajectories' level and $Z_{li}$ is the corresponding value of $l$th predictor variable at group level. $\varepsilon_{ijk}$ is the zero mean gaussian error, $\sigma_I^2$ is the component of variation owing to variability among groups, $\sigma_U^2$ is the component of variation owing to variability among pedestrian (or pair of pedestrians) nested within groups, and $\sigma_E^2$ is the component of residual variability in the repeated measurements within pair of pedestrians or pedestrians. The usual assumptions are made about the random effects: $u_{ij}$ and $b_i$ are assumed to be Gaussian distributed and uncorrelated across the levels.

The variance inflation factor (VIF) was used to verify that there was no multicollinearity among the independent variables. It was found that none of the independent variables in the models for walking speed and pedestrian

**Table 2.** Description of Independent Variables used in the Study

| Description variable | Abbreviation |
|---|---|
| External factors | |
|   Study site | Site |
|   Phase: before (B1 e B2), during (D1 e D2) and after (A1 e A2) the road crossing | Phase |
|   Effective width of zone (sidewalks or crosswalk) of each study site (m) | Width |
|   Average lane width (m) | W_Lanes |
|   Crosswalk length (m) | L_Crosswalk |
|   Presence of parking places: yes or no | Parked |
|   Road pavement: asphalt concrete or cobblestones | Pav |
|   Distance from the crosswalk to the nearest intersection (m) | Dist |
|   Approaching vehicles: yes or no | Veh |
|   Interaction with other pedestrians: yes or no | Ped |
| Social factors | |
|   Sex composition of the group of participants: female, male and mix | Sex |
|   Size group: 1, 2, or 3 pedestrians | Size |
|   Difference between the length of the legs of each pair of pedestrians | D_Leg |
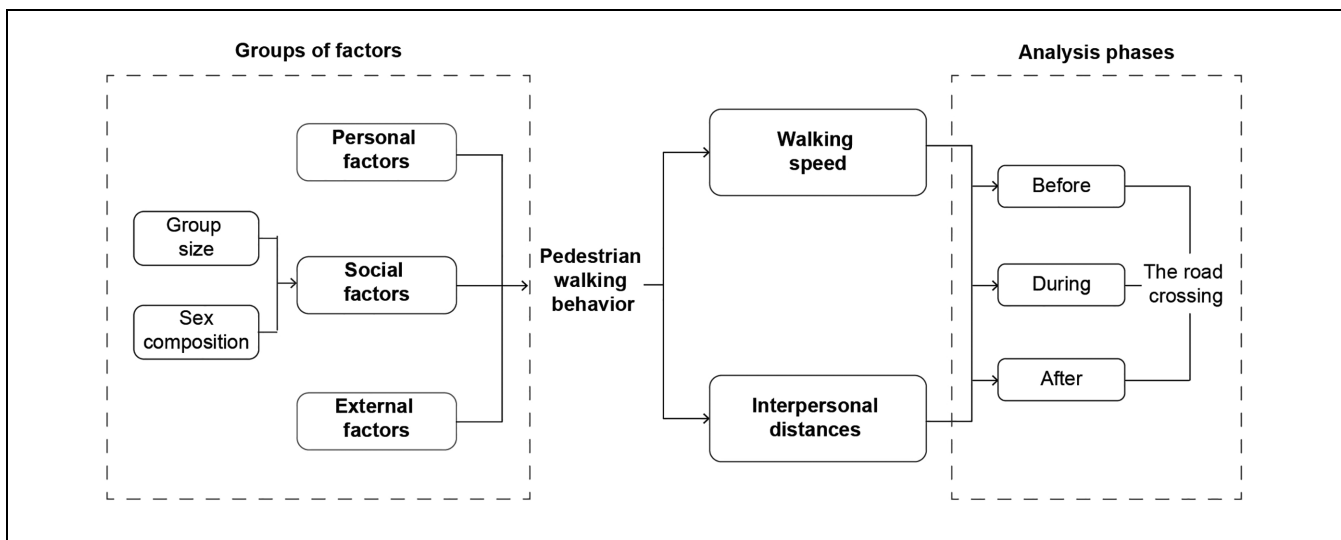


**Figure 9.** Pedestrian walking behavior analysis scheme.

distance had a VIF greater than 5 (*45*). Variance components are used to calculate the intraclass correlation (ICC), a statistic that quantifies the degree to which data at the lower level are correlated (*46*). The ICC, also referred to as the variance partition coefficient (VPC), is obtained as the proportion of random effects variance over the total variation and it ranges from zero to one.

$ICC_1$ quantifies the correlation among all the values between and within pairs of pedestrians (distance) or each pedestrian (speed) nested within groups and is given by:

$$ICC_1 = \sigma_I^2 / (\sigma_I^2 + \sigma_U^2 + \sigma_E^2) \qquad (2)$$

Similarly, $ICC_2$ quantifies the correlation among the repeated measurements within pairs of pedestrians (distance) or each pedestrian (speed) nested within groups and is given by:

$$ICC_2 = (\sigma_I^2 + \sigma_U^2) / (\sigma_I^2 + \sigma_U^2 + \sigma_E^2) \qquad (2.1)$$

This work's modelling approach followed the backward deletion method, which consisted of iteratively removing the statistically non-significant variable with the highest p-value. The final model for each response variable, namely walking speed and distance between pair of pedestrians, presents the independent variables that were statistically significant to a level of 5%. All statistical analyses were performed using the R statistical software (*47*).
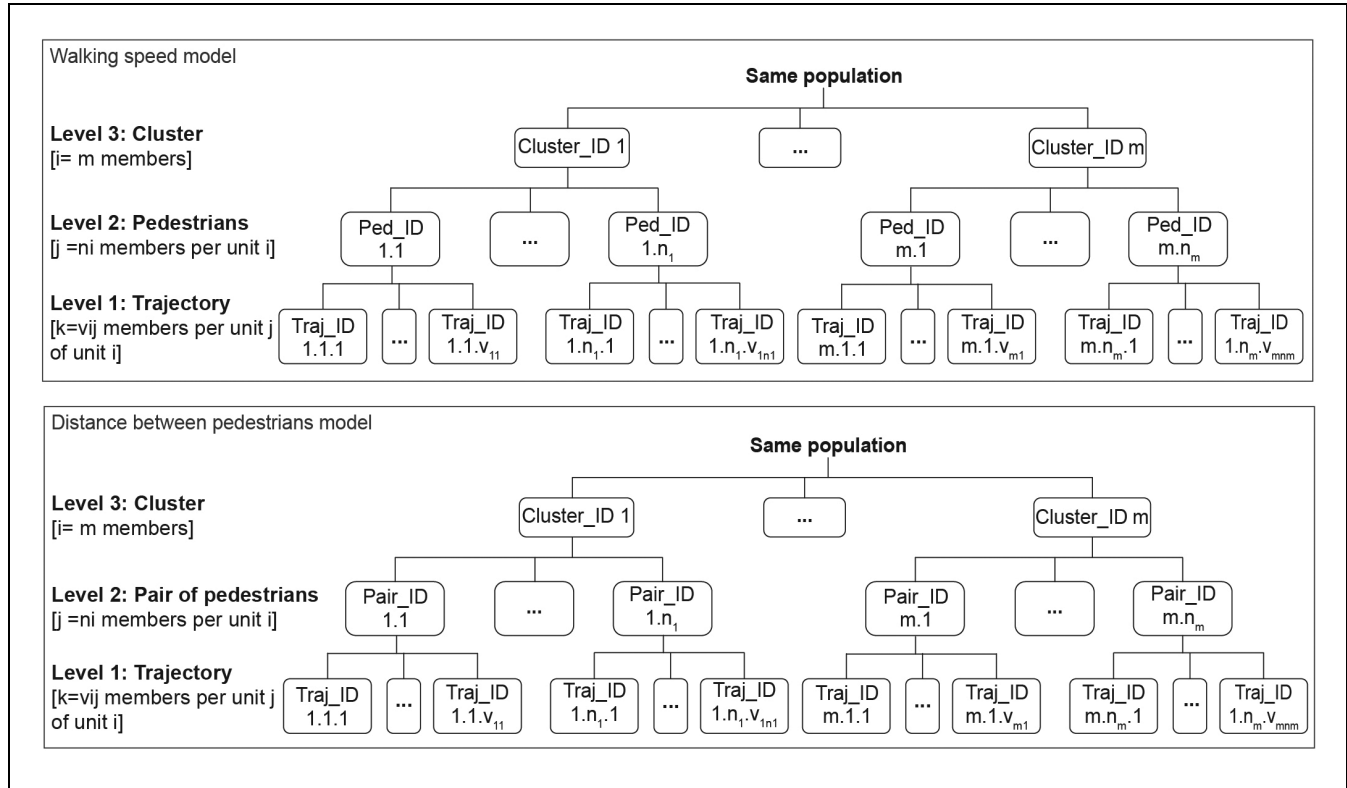
**Figure 10.** Multilevel diagram of walking speed and distance between pedestrian distance models.

## Results

To study the pedestrian behavior when crossing the road in a group, two separate models were developed for the dependent variables: (i) the average walking speed of each participant (see section 3.1), and for (ii) the distance between pairs of participants (section 3.2).

### Average Walking Speed

Traffic regulations in Portugal require that drivers cede passage to pedestrians when approaching a crosswalk without traffic lights. However, in 5% of the study's sample, drivers did not adhere to this regulation, resulting in pedestrians having to wait for the car to pass. Two trajectories with the same participants are shown in Figure 11. One can observe that in the first case (Figure 11a), the vehicle did not stop, forcing pedestrians to considerably reduce their speed compare with the second trajectory (Figure 11b). Only trajectories in which drivers complied with traffic regulations were selected for analysis (5,676 observations).

Table 3 presents summary statistics for pedestrians' walking speed, as a function of social and external factors. On average, the highest walking speed was observed for isolated individuals rather than those in groups, male

groups, in the absence of approaching vehicles, and during road crossings.

The pedestrians' walking speed was estimated using a LMM. Random effects were considered for pedestrian and group of participants, to account for repeated measurements. The final model includes all significant predictor variables (Equation 3). Non-significant social factor was $D\_Leg$ (leg length) and, external factors were $Pav$ (pavement), $W\_Lanes$ (lane width), $L\_Crosswalk$ (crosswalk length). The full summary statistics of the model are presented in Appendix B. Contrasts were used to analyze the interactions between variables and to compare the levels within each variable.

For $i = 1, \ldots, 18; j = 1, \ldots, 5, k = 1, \ldots, mij$, the model, can be written as

$$
\begin{aligned}
WS_{ijk} =\ & \beta_O + Size_{ijk} \times Phase_{ijk} + Size_{ijk} \times Ped_{ijk} \\
& + Size_{ijk} \times Park_{kij} + Sex_{ijk} \times Phase_{ijk} + Sex_{ijk} \times Ped_{ijk} \\
& + Sex_{ijk} \times Park_{ijk} + Sex_{ijk} \times Veh_{ijk} + Phase_{ijk} \\
& \times Veh_{ijk} + Phase_{ijk} \times Park_{ijk} + Veh_{ijk} \times Park_{ijk} \\
& + Width_{ijk} + Dist_i + u_{ij} + b_i + \varepsilon_{ijk}
\end{aligned}
$$

$$(3)$$

Significant effects of variables on walking speed were observed based on Satterthwaite's test (Table 4).
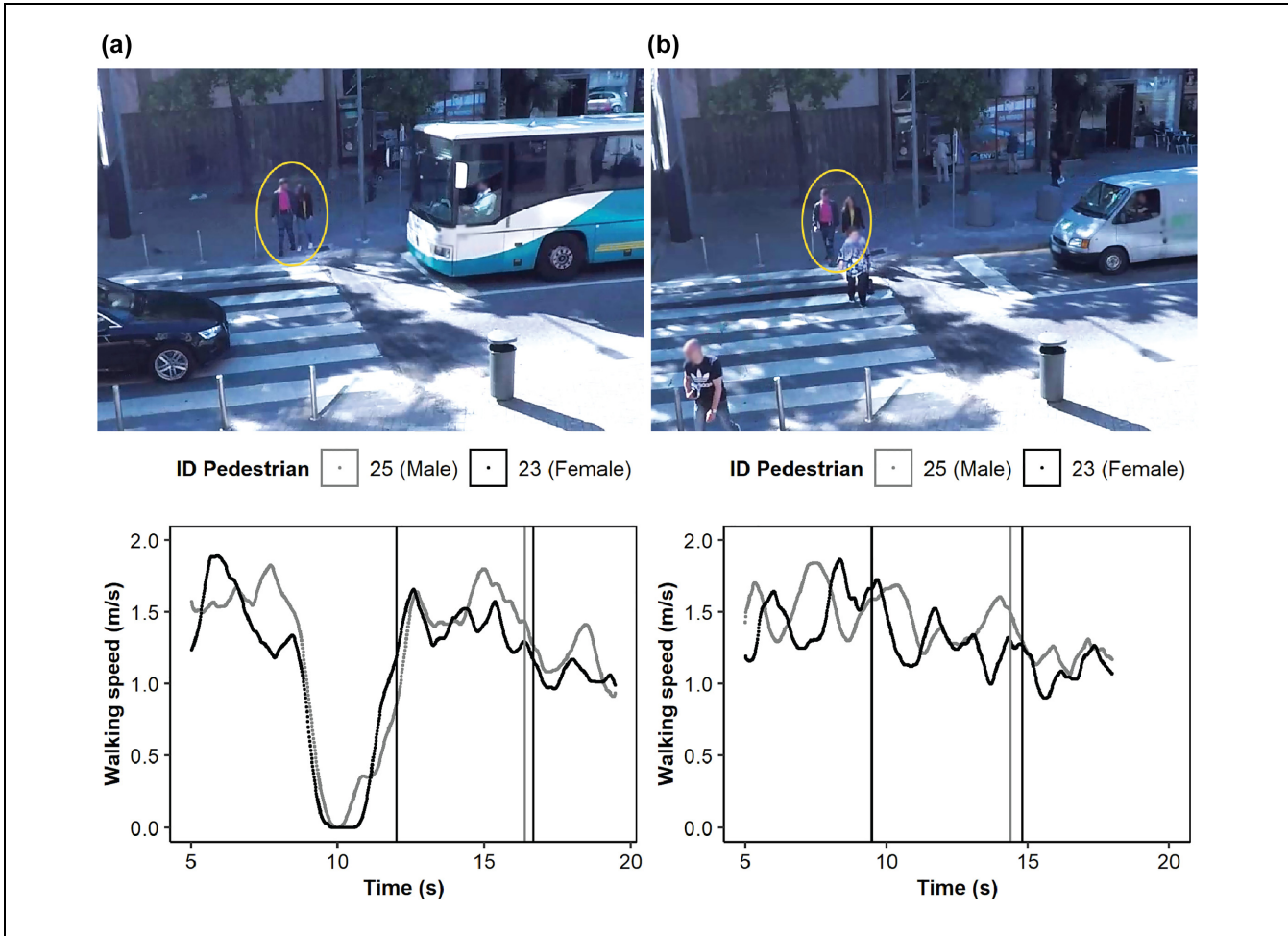
**Figure 11.** Differences between pedestrians' speed: (*a*) the vehicle did not stop, (*b*) the vehicle stopped.

Significant predictor variables and interactions are discussed below.

*Social Factors.* The statistically significant interactions between group size and the crossing phase were analyzed using contrasts. As shown in Figure 12, the pedestrians' average speed varied with the crossing phase. The average speed decreased as pedestrians approached the road crossing (B2), increased during the road crossing (D1 and D2), and then decreased as pedestrians returned to the sidewalk on the other side of the street. Notably, the average speed in B1 was similar to that in A2, suggesting that pedestrians return to their comfortable walking speed after crossing.

Figure 12 also revealed that speed depended on group size, with pedestrian speed decreasing as group size increased. The highest speed was observed for singles and the lowest for triads, and this decrease was consistent throughout the phases. Table 5 indicated that the difference between pedestrian speeds according to group size was significant (p-value < 0.05) at each crossing phase.

The interactions between variables group size (Size) with the presence of out-group pedestrians (Ped) and parked vehicles (Parked) were also significant (see Figure 13).

Table 6 presents the contrast comparing the effect of the presence of pedestrians outside the group (Ped) and parked vehicles (Parked) for each group size (Size). These results reveal that the difference between having or not having interactions with pedestrians outside the group was significant only for triads. In the case of interactions with parked vehicles, it was significant for dyads and triads. Without these interactions, the observed pedestrians' speed was higher.

The average speed also varies according to the group sex (Figure 14). It is important to emphasize that although there was a similar pattern of average speed according to the phases, the pattern of the average speed per group sex was not as homogeneous as per group size.

Concerning the phase before crossing (B1 and B2), there was also a marked reduction in the average speed

**Table 3.** Descriptive Statistics of the Walking Speed

| Variable (Abbreviation) | Walking speed (m/s) | | | |
| --- | --- | --- | --- | --- |
| | Min. | Max. | Mean | SD |
| Social factors | | | | |
| Size group (Size) | | | | |
| 1 | 0.897 | 1.978 | 1.435 | 0.184 |
| 2 | 0.660 | 2.000 | 1.318 | 0.176 |
| 3 | 0.452 | 1.922 | 1.253 | 0.197 |
| Sex composition of the group (Sex) | | | | |
| Female | 0.581 | 1.880 | 1.278 | 0.204 |
| Male | 0.683 | 1.978 | 1.329 | 0.196 |
| Mix | 0.452 | 2.000 | 1.270 | 0.191 |
| External factors | | | | |
| Study site (Site) | | | | |
| S1 | 0.581 | 2.000 | 1.289 | 0.211 |
| S2 | 0.622 | 1.857 | 1.272 | 0.178 |
| S3 | 0.452 | 1.900 | 1.293 | 0.200 |
| Crossing phase (Phase) | | | | |
| Before | | | | |
| B1 | 0.737 | 1.821 | 1.270 | 0.178 |
| B2 | 0.452 | 1.806 | 1.158 | 0.216 |
| During | | | | |
| D1 | 0.789 | 1.900 | 1.360 | 0.169 |
| D2 | 1.019 | 1.931 | 1.372 | 0.138 |
| After | | | | |
| A1 | 0.612 | 2.000 | 1.287 | 0.214 |
| A2 | 0.688 | 1.879 | 1.259 | 0.173 |
| Approaching vehicles (Veh) | | | | |
| Yes | 0.452 | 2.000 | 1.274 | 0.201 |
| No | 0.612 | 1.900 | 1.305 | 0.186 |
| Interaction with other pedestrians (Ped) | | | | |
| Yes | 0.622 | 2.000 | 1.249 | 0.200 |
| No | 0.452 | 1.978 | 1.290 | 0.195 |
| Presence of parked vehicles (Parked) | | | | |
| Yes | 0.452 | 1.978 | 1.267 | 0.198 |
| No | 0.612 | 2.000 | 1.291 | 0.196 |

*Note*: Min. = minimum; Max. = maximum; SD = standard deviation.

**Table 4.** Satterthwaite's Test Results for the Model for the Pedestrian's Average Speed

| Variables | Sum Sq | Num DF | F value | p-value |
| --- | --- | --- | --- | --- |
| Size | 6.271 | 2.000 | 145.144 | <0.001 |
| Sex | 0.507 | 2.000 | 11.729 | <0.001 |
| Phase | 12.256 | 5.000 | 113.474 | <0.001 |
| Ped | 0.039 | 1.000 | 1.824 | 0.177 |
| Parked | 0.202 | 1.000 | 9.345 | 0.002 |
| Veh | 0.428 | 1.000 | 19.817 | <0.001 |
| Width | 1.145 | 1.000 | 53.008 | <0.001 |
| Dist | 0.121 | 1.000 | 5.606 | 0.028 |
| Size × Phase | 0.457 | 10.000 | 2.118 | 0.020 |
| Size × Ped | 0.333 | 2.000 | 7.718 | <0.001 |
| Size × Parked | 0.348 | 2.000 | 8.049 | <0.001 |
| Sex × Phase | 0.838 | 10.000 | 3.880 | <0.001 |
| Sex × Ped | 0.159 | 2.000 | 3.683 | 0.025 |
| Sex × Parked | 0.205 | 2.000 | 4.736 | 0.009 |
| Sex × Veh | 0.219 | 2.000 | 5.064 | 0.006 |
| Phase × Veh | 1.288 | 5.000 | 11.928 | <0.001 |
| Phase × Parked | 0.895 | 5.000 | 8.282 | <0.001 |
| Parked × Veh | 0.081 | 1.000 | 3.737 | 0.053 |

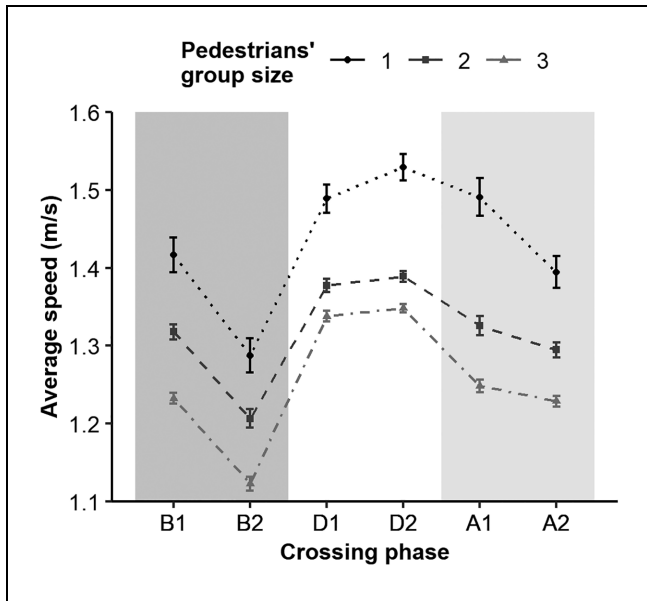*Note*: Sum Sq = sum of squares; Num DF = degrees of freedom.

**Figure 12.** Pedestrians' average speed and respective standard error according to the group size per crossing phase.

that was more pronounced for women. The reduction in speed between B1 and B2, in most cases (1, 2, and 3 pedestrians and mixed groups and all-men), was between 8% and 9%, except for female groups where the average speed reduction was of 13% (Table 7— phase B2). Additionally, between B2 and D1, the average speed increase in the female groups was also the highest (24%). Although this study did not evaluate the decision-making process to cross, these results indicate that the group sex could be a relevant factor for the crossing time.

As can be noted in Table 8, there was a clear difference between the average speed of the all-male groups compared with the all-female ones before and after the crossing. The differences between all-female and mixed groups were only significant before the crossing. This could indicate that women tend to slow down even more when deciding to cross, which is why only in this phase there were significant differences from the mixed groups.

Interactions between group sex (Sex) and the presence of out-group pedestrians (Ped), parked (Parked) and approaching vehicles (Veh) were also found to be significant. The results of the contrast between these variables are shown in Table 9.

As indicated by the model (see also Figure 15), the speed increase in the female groups was significant when there were no out-group pedestrians compared with when there were (Figure 15a). In the male groups, the increase in speed was significant when there were no parked vehicles versus when there were parked vehicles (Figure 15b). Finally, in these two types of groups (male or female), the increase in speed was significant when no vehicle was approaching (Figure 15c).

*External Factors.* With reference to the external factors influencing pedestrian walking speed, the interaction between approaching vehicles (Veh) and parked vehicles (Parked) was significant, as were the interactions between approaching (Veh) and parked vehicles (Parked) with the crossing phase (Phase).

On the one hand, the model results show that differences in the pedestrians' speed with or without parked vehicles were significant only when the vehicles approached ($\hat{\beta} = -0.031$, SE = 0.008, p-value < 0.001). When there were no approaching vehicles, the effect of parked vehicles was not significant ($\hat{\beta} = -0.011$, SE = 0.009, p-value = 0.229) (see Figure16).

On the other hand, in the interactions between approaching and parked vehicles with the crossing phase, as shown in Figure 17, a and b, there was a similar pattern in both cases, according to the crossing phase. This pattern was similar to the one mentioned previously: speed reduction before the crossing, speed increase during the crossing, and decrease again after the crossing. The results of the contrasts used to analyze these interactions are presented in Table 10.

**Table 5.** Group Size Contrasts According to the Phase.

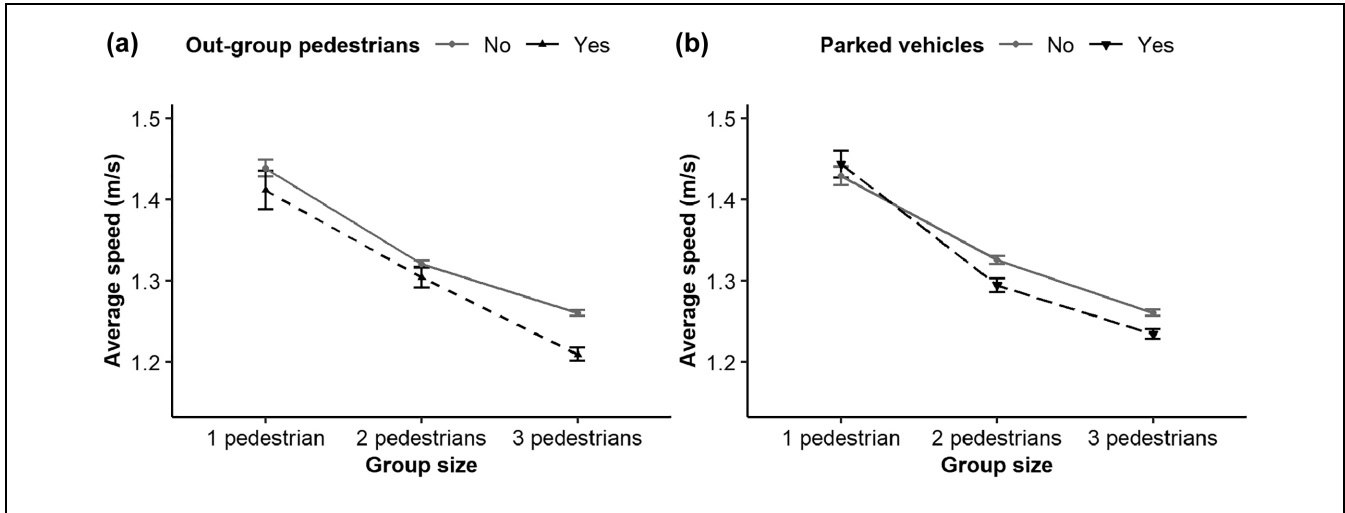| | Size contrasts | | | | | | | | |
| | 1-2 | | | 1-3 | | | 2-3 | | |
| Phase | Estimate | SE | p-value | Estimate | SE | p-value | Estimate | SE | p-value |
|---|---|---|---|---|---|---|---|---|---|
| B1 | 0.088 | 0.023 | <0.001 | 0.182 | 0.023 | <0.001 | 0.094 | 0.012 | <0.001 |
| B2 | 0.093 | 0.023 | <0.001 | 0.182 | 0.022 | <0.001 | 0.089 | 0.012 | <0.001 |
| D1 | 0.104 | 0.024 | <0.001 | 0.158 | 0.023 | <0.001 | 0.054 | 0.013 | <0.001 |
| D2 | 0.126 | 0.024 | <0.001 | 0.190 | 0.023 | <0.001 | 0.063 | 0.013 | <0.001 |
| A1 | 0.154 | 0.023 | <0.001 | 0.247 | 0.022 | <0.001 | 0.093 | 0.012 | <0.001 |
| A2 | 0.090 | 0.024 | <0.001 | 0.170 | 0.023 | <0.001 | 0.080 | 0.012 | <0.001 |

*Note:* SE = standard error.

**Figure 13.** Pedestrians' average speed and respective standard error per group size according to (*a*) presence of out-group pedestrians, and (*b*) presence of parked vehicles.

**Table 6.** Contrast of Interactions with Out-Group Pedestrians and Parked Vehicles per Group Size

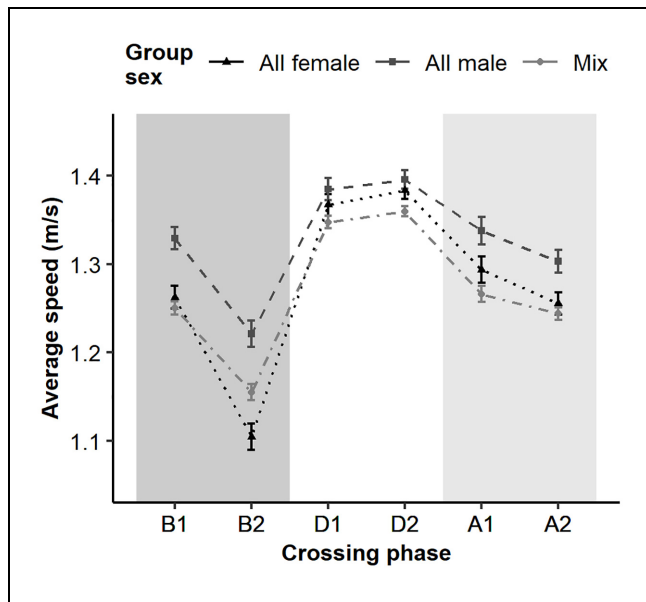| | Ped | | | Parked | | |
|---|---|---|---|---|---|---|
| | No—Yes | | | No—Yes | | |
| Size | Estimate | SE | p-value | Estimate | SE | p-value |
| 1 pedestrian | 0.017 | 0.021 | 0.437 | −0.027 | 0.017 | 0.102 |
| 2 pedestrians | −0.018 | 0.011 | 0.122 | 0.044 | 0.009 | <0.001 |
| 3 pedestrians | 0.035 | 0.008 | <0.001 | 0.045 | 0.007 | <0.001 |

*Note*: SE = standard error.



**Figure 14.** Pedestrians' average speed and respective standard error according to the group sex per crossing phase.

Without approaching vehicles, the average speed was higher in phases B1, B2, D1, and A2 (Figure 17*a*), and the model confirm a significant positive difference. On the contrary, there was a significant difference with a negative coefficient in the first segment after the crossing. These results seem to indicate that when pedestrians confirm that there are no approaching vehicles, they begin to slow down during the road crossing (Phase D2), as on reaching the other side of the road (phase A1), their speed was lower in comparison to when vehicles were approaching. Conversely, when a vehicle approached or waited, pedestrians increased their speed during the crossing.

The pattern according to the phases without parked vehicles was similar to that without approaching vehicles (Figure 17*b*). Although smaller differences between the speed with or without a parked vehicle exist before the crossing, the contrasts confirm these positive differences.

Before the road crossing (B1 and B2) if there was a parked vehicle, the model results showed that the speed of pedestrians decreases, perhaps owing to difficulty of seeing the oncoming vehicles. During the crossing in D2, the pedestrians increased their speed even more if there

**Table 7.** Pedestrian Average Speed by the Phase According to the Size and to the Sex of the Group

| Phase | Size | Average speed (m/s) | SE | Speed variation | Sex | Average speed (m/s) | SE | Speed variation |
|---|---|---|---|---|---|---|---|---|
| B1 | 1 | 1.417 | 0.022 | na | Female | 1.263 | 0.013 | na |
| B2 | 1 | 1.287 | 0.022 | −9% | Female | 1.102 | 0.015 | −13% |
| D1 | 1 | 1.489 | 0.018 | 16% | Female | 1.366 | 0.012 | 24% |
| D2 | 1 | 1.529 | 0.017 | 3% | Female | 1.384 | 0.010 | 1% |
| A1 | 1 | 1.491 | 0.024 | −2% | Female | 1.294 | 0.015 | −6% |
| A2 | 1 | 1.395 | 0.021 | −6% | Female | 1.256 | 0.012 | −3% |
| B1 | 2 | 1.318 | 0.010 | na | Male | 1.329 | 0.012 | na |
| B2 | 2 | 1.206 | 0.012 | −8% | Male | 1.219 | 0.015 | −8% |
| D1 | 2 | 1.377 | 0.009 | 14% | Male | 1.385 | 0.012 | 14% |
| D2 | 2 | 1.389 | 0.007 | 1% | Male | 1.397 | 0.011 | 1% |
| A1 | 2 | 1.326 | 0.012 | −5% | Male | 1.338 | 0.016 | −4% |
| A2 | 2 | 1.295 | 0.010 | −2% | Male | 1.304 | 0.013 | −3% |
| B1 | 3 | 1.232 | 0.007 | na | Mix | 1.250 | 0.007 | na |
| B2 | 3 | 1.123 | 0.009 | −9% | Mix | 1.147 | 0.009 | −8% |
| D1 | 3 | 1.338 | 0.007 | 19% | Mix | 1.345 | 0.007 | 17% |
| D2 | 3 | 1.348 | 0.006 | 1% | Mix | 1.359 | 0.006 | 1% |
| A1 | 3 | 1.248 | 0.008 | −7% | Mix | 1.266 | 0.009 | −7% |
| A2 | 3 | 1.229 | 0.007 | −2% | Mix | 1.245 | 0.007 | −2% |

*Note:* SE = standard error; na = not applicable (the variation in speed is determined calculating the difference with the previous segment of phase).

**Table 8.** Contrasts of the Group Sex Per Crossing Phase

| | Sex contrasts | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Female—Male | | | Female—Mix | | | Male—Mix | | |
| Phase | Estimate | SE | p-value | Estimate | SE | p-value | Estimate | SE | p-value |
| B1 | −0.077 | 0.019 | 0.000 | −0.023 | 0.016 | 0.307 | 0.054 | 0.015 | 0.001 |
| B2 | −0.119 | 0.018 | <0.001 | −0.080 | 0.015 | <0.001 | 0.038 | 0.015 | 0.025 |
| D1 | −0.026 | 0.019 | 0.343 | −0.008 | 0.016 | 0.875 | 0.018 | 0.015 | 0.457 |
| D2 | −0.018 | 0.019 | 0.587 | −0.008 | 0.016 | 0.863 | 0.010 | 0.015 | 0.782 |
| A1 | −0.052 | 0.018 | 0.011 | −0.014 | 0.015 | 0.626 | 0.038 | 0.015 | 0.026 |
| A2 | −0.059 | 0.019 | 0.004 | −0.019 | 0.016 | 0.450 | 0.040 | 0.015 | 0.020 |

*Note:* SE = standard error.

**Table 9.** Contrast of Interactions with Out-Group Pedestrians, Parked Vehicles and Approaching Vehicles Per Group Sex

| | Ped | | | Parked | | | Veh | | |
|---|---|---|---|---|---|---|---|---|---|
| | No—Yes | | | No—Yes | | | No—Yes | | |
| Sex | Estimate | SE | p-value | Estimate | SE | p-value | Estimate | SE | p-value |
| Female | 0.034 | 0.013 | 0.009 | 0.011 | 0.011 | 0.309 | 0.024 | 0.011 | 0.032 |
| Mix | 0.011 | 0.012 | 0.330 | 0.007 | 0.009 | 0.434 | 0.010 | 0.007 | 0.131 |
| Male | −0.012 | 0.013 | 0.343 | 0.044 | 0.011 | <0.001 | 0.046 | 0.010 | <0.001 |

*Note:* SE = standard error.

was a parked vehicle. This behavior could be explained by the pedestrians' desire to get closer to the parked vehicle faster as a way of becoming protected from an approaching vehicle. As previously shown, the difference between having or not parked vehicles was only significant when vehicles were on approach. After the crossing, the speed was lower, and the differences were significant.

Finally, within the characteristics of the infrastructure, there were significant effects of the effective width of the
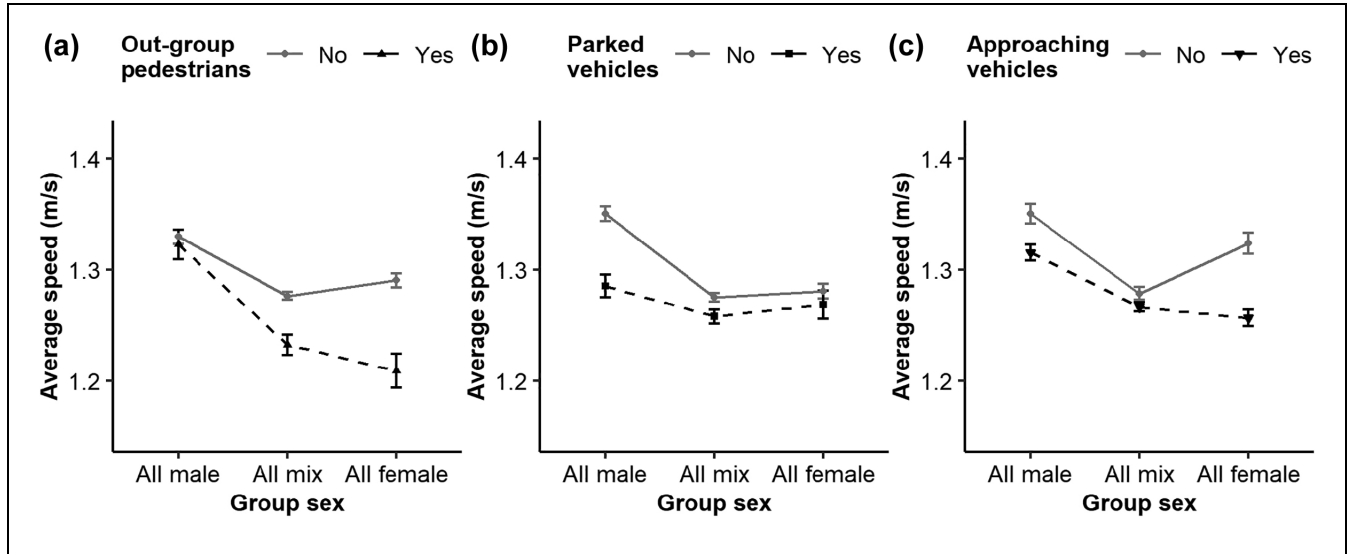
**Figure 15.** Pedestrians' average speed and respective standard error per group sex according to (*a*) presence of out-group pedestrians, (*b*) presence of parked vehicles, and (*c*) presence of approaching vehicles.
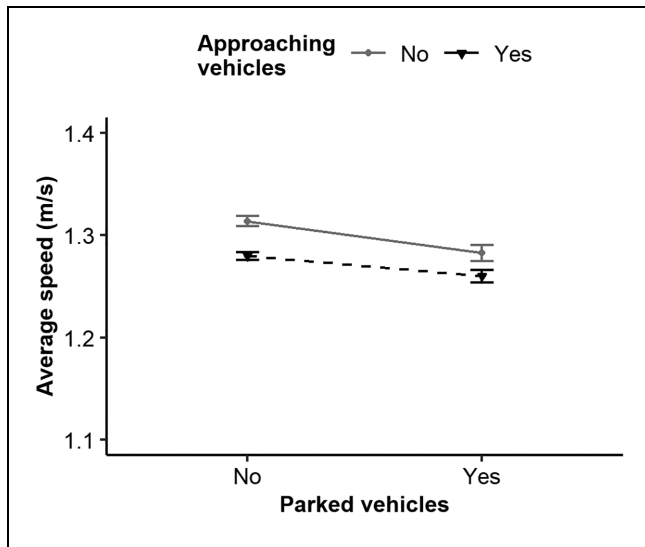


**Figure 16.** Pedestrians' average speed and respective standard error according to the presence of approaching vehicles per presence of parking vehicles.

zone, $\hat{\beta} = -0.013$, SE = 0.002, p-value 0.001 (wider pedestrian zones were associated with lower speeds), and the distance between the pedestrian crossing and the nearest intersection, $\hat{\beta} = 0.003$, SE = 0.001, p = 0.02 (the greater the distance, the lower the speed of the pedestrians).

## Average Distance Between Pair of Pedestrians

A sample of 3279 observations of the distance between pairs of pedestrians was analyzed. As explained in section 2.3, the random effects considered for these models were the ID of the pair of participants and the ID of the group of participants.

Table 11 shows summary statistics for the average distance between pairs of pedestrians according to the same variables analyzed the walking speed. With these results, it was possible to observe that, on average, the distance between pairs of pedestrians was greater in triads compared with dyads, in male groups compared with female and mixed, and during crossing compared with before.

As with the pedestrians' walking speed, the model was fitted using LMMs. The final model contains all significant predictor variables (Equation 4). Full summary statistics of the model are presented in Appendix C. Contrasts were used to analyze the interactions between variables and to summarize the model results according to each variable.

For $i = 1, \ldots, 18; j = 1, \ldots, 5, k = 1, \ldots, mij$, the model, can be written as:

$$D_{ijk} = \beta_O + Size_{ijk} \times Sex_{ijk} + Size_{ijk} \times Park_{kij} + Phase_{ijk} + Veh_{ijk} + Width_{ijk} + u_{ij} + b_i + \varepsilon_{ijk} \qquad (4)$$

Satterthwaite's test was used to identify variables with significant influence on the average distance between pedestrians, and the interactions. The results are presented in Table 12. The non-significant social factor was *D_Leg* (leg length) and external factors were: *Pav* (pavement), *W_Lanes* (lane width), *L_Crosswalk* (crosswalk length), *Dist* (Distance from the crosswalk to the nearest intersection), *and Ped* (interaction with other pedestrians).
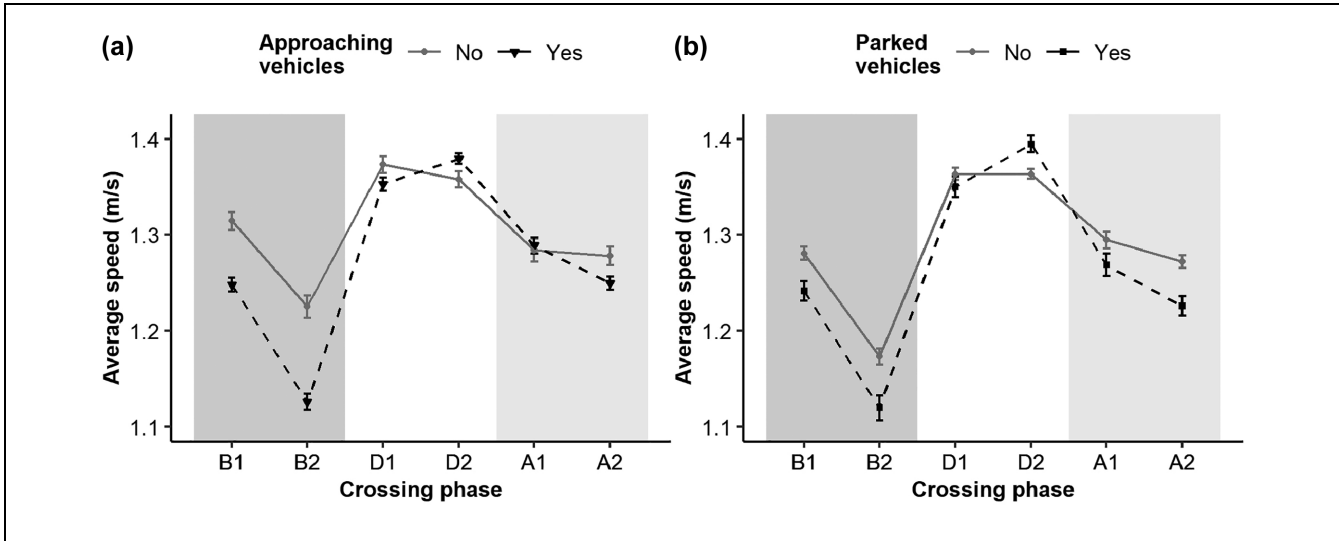
**Figure 17.** Pedestrians' average speed and respective standard error per crossing phase according to (*a*) presence of approaching vehicles, and (*b*) presence of parked vehicles.

**Table 10.** Approaching Vehicles and Parked Vehicles Contrasts According to the Phase

| | Veh | | | Parked | | |
|---|---|---|---|---|---|---|
| | No—Yes | | | No—Yes | | |
| Phase | Estimate | SE | p-value | Estimate | SE | p-value |
| B1 | 0.051 | 0.011 | <0.001 | 0.032 | 0.012 | 0.008 |
| B2 | 0.077 | 0.011 | <0.001 | 0.050 | 0.012 | <0.001 |
| D1 | 0.037 | 0.012 | 0.002 | 0.010 | 0.012 | 0.402 |
| D2 | −0.002 | 0.012 | 0.866 | −0.036 | 0.012 | 0.003 |
| A1 | −0.022 | 0.011 | 0.048 | 0.025 | 0.012 | 0.034 |
| A2 | 0.022 | 0.011 | 0.054 | 0.044 | 0.012 | <0.001 |

*Note:* SE = standard error.

*Social Factors.* The interactions between group size and sex were statistically significant. As illustrated in Figure 18, distances between pedestrians were greater for triads than for dyads. The variation was 10% for all-female, 7% for all-male, and 4% for mixed groups.

Contrasts were utilized to analyze the effect of the interaction between sex composition (Sex) and group size (Size) on the distance between pedestrians. When comparing groups of dyads and triads based on sex, significant differences were observed in male groups ($\hat{\beta}$ = −0.053, SE = 0.022 p = 0.016) and female groups ($\hat{\beta}$ = −0.104, SE = 0.024, p-value < 0.001), but not in mixed groups ($\hat{\beta}$ = −0.017, SE = 0.014, p = 0.056).

Furthermore, male pairs appear to maintain a greater distance between them compared with mixed or female pairs. The contrasts for the interaction between sex and size revealed a significant difference between males and females for dyads and triads. There was only a difference

between male and mixed groups for triads, as shown in Table 13.

Interactions of group size (Size) with parked vehicles (Parked) were also significant for the distance between pedestrians. As already mentioned, the distance is greater when the size of the group increases. However, Figure 19 shows that without parked vehicles, the distance was greater than with parked vehicles, particularly for groups of two pedestrians. The contrast confirms these positive differences. ($\hat{\beta}$ = 0.050, SE = 0.019, p = 0.009). The differences in triads was not significant ($\hat{\beta}$ = −0.004, SE = 0.011, p = 0.745).

*External Factors.* Concerning external factors, the effective width of the zone (Width) ($\hat{\beta}$ = −0.010, SE = 0.003, p = 0.002) had a significant effect on the average distance. The distance between pedestrians decreases with

**Table 11.** Descriptive Statistics of the Distance Between Pairs of Pedestrians

| Variable (Abbreviation) | Distance between pair of pedestrians (m) | | | |
| --- | --- | --- | --- | --- |
| | Min. | Max. | Mean | SD |
| Social factors | | | | |
|   Size group (Size) | | | | |
|     1 | na | na | na | na |
|     2 | 0.427 | 2.129 | 0.813 | 0.192 |
|     3 | 0.376 | 2.254 | 0.866 | 0.258 |
|   Sex composition of the group of participants (Sex) | | | | |
|     Female | 0.388 | 2.101 | 0.840 | 0.226 |
|     Male | 0.431 | 2.254 | 0.894 | 0.260 |
|     Mix | 0.376 | 2.237 | 0.838 | 0.242 |
| External factors | | | | |
|   Study site (Site) | | | | |
|     S1 | 0.376 | 1.729 | 0.821 | 0.204 |
|     S2 | 0.430 | 2.129 | 0.852 | 0.258 |
|     S3 | 0.385 | 2.254 | 0.884 | 0.261 |
|   Crossing phase (Phase) | | | | |
|     Before | | | | |
|       B1 | 0.413 | 1.699 | 0.787 | 0.210 |
|       B2 | 0.425 | 2.047 | 0.860 | 0.242 |
|     During | | | | |
|       D1 | 0.477 | 2.251 | 0.902 | 0.234 |
|       D2 | 0.475 | 2.237 | 0.862 | 0.220 |
|     After | | | | |
|       A1 | 0.376 | 2.254 | 0.852 | 0.265 |
|       A2 | 0.385 | 2.114 | 0.852 | 0.271 |
|   Approaching vehicles (Veh) | | | | |
|     Yes | 0.376 | 2.129 | 0.854 | 0.242 |
|     No | 0.385 | 2.254 | 0.849 | 0.247 |
|   Interaction with other pedestrians (Ped) | | | | |
|     Yes | 0.413 | 1.785 | 0.864 | 0.244 |
|     No | 0.376 | 2.254 | 0.850 | 0.244 |
|   Presence of parked vehicles (Parked) | | | | |
|     Yes | 0.385 | 1.872 | 0.838 | 0.244 |
|     No | 0.376 | 2.254 | 0.858 | 0.244 |

*Note*: Min. = minimum; Max. = maximum; SD = standard deviation; na = not applicable.

**Table 12.** Satterthwaite's Test Results for the Model of the Average Distance Between Pairs of Pedestrians

| Variable | Sum Sq | Num DF | F value | p-value |
| --- | --- | --- | --- | --- |
| Sex | 0.473 | 2.000 | 4.870 | 0.010 |
| Size | 1.175 | 1.000 | 24.178 | 0.000 |
| Phase | 6.161 | 5.000 | 25.353 | <0.001 |
| Veh | 0.383 | 1.000 | 7.888 | 0.005 |
| Parked | 0.202 | 1.000 | 4.151 | 0.042 |
| Width | 0.490 | 1.000 | 10.073 | 0.002 |
| Sex × Size | 0.400 | 2.000 | 4.113 | 0.016 |
| Size × Parked | 0.299 | 1.000 | 6.142 | 0.013 |

*Note*: Sum Sq = sum of squares; Num DF = degrees of freedom.

wider pedestrian zones. These results are contrary to what was expected. With less space, shorter distances between people were the expectation.

Furthermore, the distance between pedestrians depends on the crossing phase (Figure 20). Near the crosswalk (B2) the distance increases and the differences between the segments (B1 and B2) before the road crossing were significant ($\hat{\beta} = -0.085$, SE = 0.013, p-value < 0.001). In addition, the distance between pedestrians was greater when a vehicle approached (Veh) ($\hat{\beta} = 0.029$,

**Table 13.** Contrasts of the Group Size Per Group Sex

| | Size | | | | | |
| | 2 pedestrians | | | 3 pedestrians | | |
| Sex | Estimate | SE | p-value | Estimate | SE | p-value |
|---|---|---|---|---|---|---|
| Female—Male | −0.144 | 0.046 | 0.005 | −0.093 | 0.038 | 0.043 |
| Female—Mix | −0.088 | 0.038 | 0.056 | −0.012 | 0.031 | 0.919 |
| Male—Mix | 0.057 | 0.038 | 0.294 | 0.081 | 0.032 | 0.032 |

*Note*: SE = standard error.
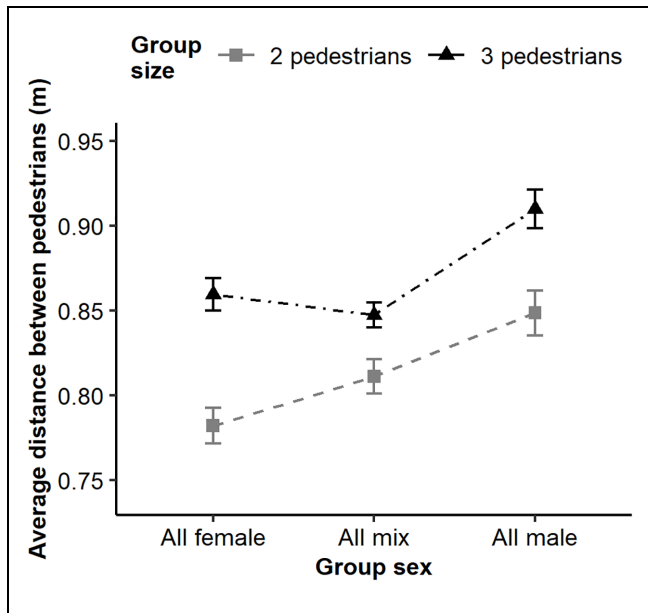


**Figure 18.** Average distance between pairs of pedestrians and respective standard error according to the group size per group sex.



**Figure 19.** Average distance between pairs of pedestrians and respective standard error according to the presence of parked vehicles per group size.

SE = 0.010, p = 0.005). This could indicate differences in the decision-making process to cross, which would increase the distance if one of the people in the pair decides to cross first.

Between the segment during (D1-D2) and after (A1-A2) the crossing there were no significant differences ($\hat{\beta} = -0.033$, SE = 0.014, p = 0.168 and $\hat{\beta} = -0.005$, SE = 0.013, p = 0.999).
During the first segment of the road crossing (D1) the average distance increased, and the difference between the previous segment (B2) was significant ($\hat{\beta} = -0.069$, SE = 0.014, p-value < 0.001). Then, in the second segment during the crossing (D2) the average distance decreased and the difference with the next segment (A1) was significant ($\hat{\beta} = 0.050$, SE = 0.015, p = 0.012). This finding suggests that as pedestrians approach the stree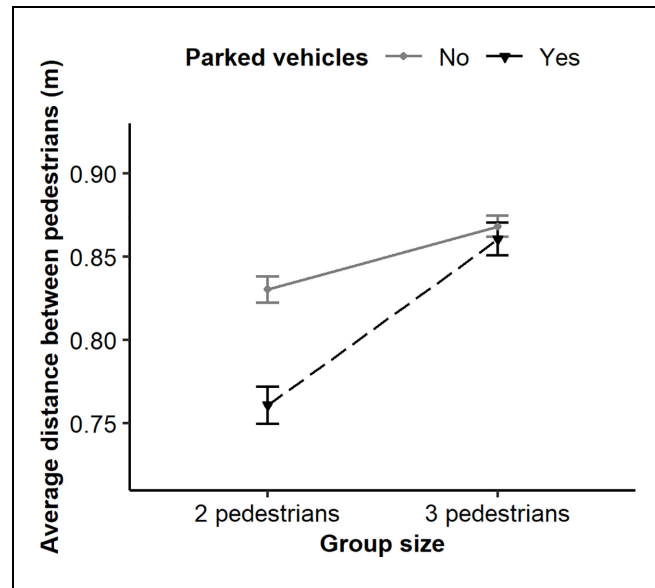t crossing area where they assess oncoming vehicles, they tend to initially separate but eventually rejoin near the end of the crossing.

## Discussion

It is indicated by the findings reported here that the dependence of walking behavior of social group pedestrians on group size and the composition by sex is also commonly observed in the road crossing context under normal conditions of low pedestrian density.

### Size

An increase in group size was found to be associated with a decrease in average speed and an increase in average distance between pairs of pedestrians. Previous studies have shown that in areas without vehicle interactions,
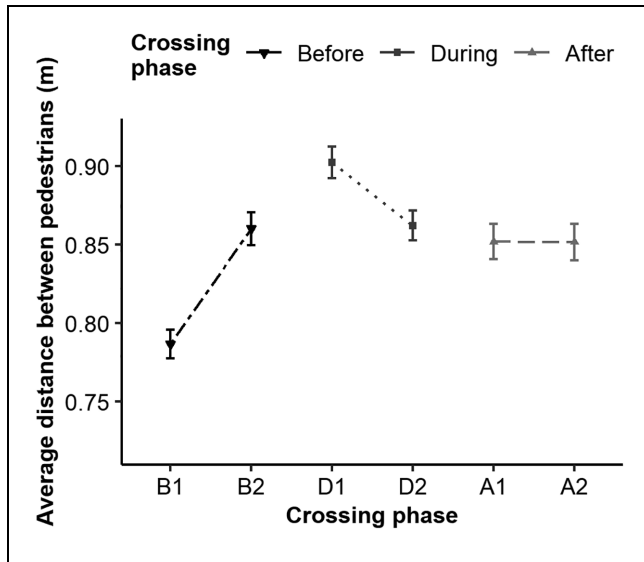
**Figure 20.** Average distance between pedestrians by crossing phase.

such as commercial areas, streets, or inside universities, social groups walk at a slower speed than isolated pedestrians (*3, 17, 30*). The decrease in speed by the groups was attributed to the need to maintain cohesion between individuals to facilitate social interactions (*3, 17*). However, these studies were conducted under medium to high pedestrian densities, and the decrease in speed could be an effect of the group's difficulty to stay together. Similarly, it was found in the present study that pedestrian groups also walked at a lower speed under low-density conditions, even while crossing the street.

With regard to the distance between pedestrians, the largest distances were observed among triads compared with dyads (*3, 27*). This could be explained by a tendency for greater intimacy among couples among triplets (*27*). Nevertheless, it is also possible that triads organize themselves differently to facilitate communication. For instance, on sidewalks, where there is limited space, they may need to adopt different formations, such as a "V"-like shape observed by Moussaïd et al. (*3*), where the distances between people are greater.

## Sex

The average speed and distance are also affected by the sex composition of the group. Firstly, a clear difference in average speed was observed between male groups and female/mixed groups before and after the road crossing, with male groups showing higher speed. These findings were expected, as males are generally associated with higher average speeds than females (*48*), and other studies have shown that male groups walk faster than females and mixed groups (*7, 27*). Costa (*7*) suggested that the

deeper the affective involvement between a couple (mixed dyads), the slower the walking speed tends to be. In this study, the participants were friends, and little difference was found between the walking speed of female and mixed dyads (except before the road crossing).

In contrast, Federici et al. (*30*) did not observe differences in walking speed according to the sex of the group. However, the analysis conditions of these two works were different, as Costa (*7*) observed the group behavior of pedestrians while walking on sidewalks and in pedestrian areas in different urban settings, while Federici et al. (*30*) focused on investigating pedestrian dynamics in medium-high density situations with observations of incoming pedestrian flows to a university admission test. This difference could indicate that in normal or/and low-density conditions, sex can be a relevant factor in the speed of pedestrians.

Secondly, male groups maintained a greater distance than the female groups (dyads and triads), as well as the mixed triads. No difference was observed between female and mixed groups. Similar results were reported by Costa (*7*), who found that male dyads and triads tended to walk together less often than female dyads, and mixed dyads walked together more often than same-sex dyads. This is consistent with the literature on the topic of interpersonal distance, which indicates that female pairs tend to maintain closer distances than male pairs, possibly owing to a greater predisposition of women to be affiliative (see Uzzell & Horne [*49*] for a review and discussion).

## Phases

It was found that neither the pedestrians' average speed nor the distance between pairs remained constant throughout the three crossing phases, according to the results. Empirical observations in previous studies of stairs inside a university had shown that the average distance for groups of different sizes is stable for the purpose of maintaining group structure (*50*). However, the results of this study showed that the distance between pairs of pedestrians changed according to the phases of the road crossing. One possible reason for this may be that individuals increase their distance from each other (separate) near the crosswalk when deciding whether to cross or not. Furthermore, since the width of the crosswalk is usually greater than the effective width of the sidewalks, the distance between pedestrians could also be expected to increase during the road crossing. However, in this study, the effect of the effective width showed that the distance was smaller in the broader zones.

With regards to the average speed, there was a significant reduction before the road crossing, followed by an increase during the crossing and a posterior reduction back to the initial speed. This pattern was expected since

before the crossing in segment B2, near the crosswalk, pedestrians evaluate whether they can cross or not, and while making that judgment, they reduce their speed. During road crossings, an increase in speed was also expected since previous studies have reported that pedestrian speeds are higher when crossing streets than on sidewalks (although the behavior of social groups has not been previously analyzed (*51*).

An important conclusion of this study is that the average speed pattern is maintained through the phases by group size (higher speed for singles and lower for groups in all phases). However, there were some differences in the pattern throughout the phases according to the groups' sex. Before the road crossing, the speed reduction was higher for female groups, followed by mixed and male groups. It was the only phase where there was a significant difference between females and mixed groups. Although the objective of this study was not to analyze the effect of sex on the decision-making process to cross, the observed differences may indicate that sex plays an important role when pedestrians decide to cross. Also, the effect of oncoming vehicles was significant on average speed by sex composition but not by group size.

This observation is consistent with those of Holland and Hill (*52*). They compared the intention to cross and the perceived risk through the answers to a questionnaire in which people were asked to evaluate whether to cross or not in hypothetical textual risk descriptions, concluding that women were less likely than men to try to cross under risky situations. Also, Díaz (*53*) analyzed the attitude toward (illegal) mid-block crossing and found that men reported more violations than women.

Unlike before and after the crossing, during the road crossing, there was no significant difference in average speed between the group sex composition. This could indicate that the behavioral approach was different between phases. During the crossing, the objective of reaching the other side as quickly as possible may override the behavioral characteristics on the sidewalks. The greater the risk involved in walking, the higher the walking speed (*51*). This contrasts the conclusion of Montufar et al. (*48*), who found significant differences between the speed of men compared with women during street crossings, although they did not study the behavior of pedestrians in groups.

### Interactions with Parked Vehicles and Out-Group Pedestrians

The results revealed that the presence of parked vehicles had an effect according to the phase. If there was a parked vehicle before the road crossing, the pedestrians' average speed decreased. On the other hand, during the crossing, pedestrians increased their speed even more.

These results may align with the conclusions of previous studies, which analyzed the acceptance of the gap for mid-block street crossings and red-light violations. On the one hand, the presence of illegal parking seems to discourage pedestrians from crossing because it affects pedestrians' visibility, forcing pedestrians to be more careful and reject shorter time gaps for crossing, according to Yannis et al. (*51*). This could be related to the decreased speed before crossing. On the other hand, according to Dommes et al. (*52*), the presence of parked vehicles generated red light violations, suggesting that parked vehicles may increase the pedestrians' feeling of safety once the road width decreases. This could be related to increased speed during the crossing once the decision to cross has been taken.

Finally, the results show that, in conditions of low pedestrian density, the differences in average speed owing to interaction with other pedestrians were significant for triads and groups of women. Groups of pedestrians slow down, perhaps to avoid splitting up. Additionally, no significant differences were found in the average distance. This may indicate that in normal situations, young pedestrians only change their speed and do not change their organization. This was not unexpected since it is known that the presence of other pedestrians strongly influences average walking speeds, depending on pedestrian density (*51*).

## Conclusions

Even though it had been revealed in previous studies that pedestrian traffic comprises groups and that social interactions are crucial for organizing human crowds, little or no research had been conducted on the behavior of pedestrians' social groups under normal conditions near pedestrian crossing sites.

This study's empirical observations confirmed that at least 40% of observed pedestrians walk in a group. The findings suggest several critical factors in the movement preferences of social groups, even under normal conditions and in places with low pedestrian density. It was found that typical walking patterns arose from local interactions between group members. Differences in speed and distance between young pedestrians depended on social factors (group size and sex composition) and external factors (crossing phases and interaction with other pedestrians, approaching and parked vehicles).

The results of this study should help to increase understanding of how young pedestrians in social groups interact with each other and with the environment, thus improving the way scientists, engineers, policymakers, and designers increase the safety of vulnerable road users. The description of several fundamental elements of

movement behavior (speed and distance) can be used to assign realistic values to models of pedestrian behavior based on parameters such as group size, sex composition, and crossing phases.

Furthermore, these findings could be used in risk studies of vehicle–pedestrian interactions and in new challenges concerning pedestrian dynamics, for example, to create rules and improve shared space areas, as well as to give indications to improve the automatic detection of pedestrians by autonomous vehicles.

As with any experiment, there are limitations and extensions that need to be further investigated in future studies. The conclusions drawn in this work are based on young voluntary participants. However, previous studies have reported that age affects the speed and distance between pedestrians, with older pedestrians walking at slower speeds and preferring greater interpersonal distances (*54*, *55*). Therefore, it is crucial to include participants from different age groups in future work on social group behavior. Additionally, future studies should consider mixed traffic conditions, including particular types of pedestrians, such as people with disabilities (i.e., wheelchair users). These individuals, for example, have different physical dimensions (i.e., weight, width, or height) that affect the required sight distance and speed, whether on sidewalks or at crosswalks (*56*).

Furthermore, conclusions drawn from these results about infrastructure factors for urban planning should be taken cautiously. The results have been obtained from experiences in a few urban settings, which may not represent a diversity of environments. Nevertheless, this work reaffirms the need to continue studying how pedestrians perceive and interpret differences in urban environments, such as the presence of obstacles. Another question that remains open is whether the organization of social groups results in more risky crossing decisions.

The coexistence of pedestrians and micro-mobility vehicles (MMV) presents a recent challenge for efficient and safe use of road infrastructure. MMVs are becoming more popular and widely used, potentially owing to the time savings they offer for short distances to their final destinations (*57*, *58*). MMVs can move at higher speeds, which can affect pedestrians' risk perception. For instance, cyclists and electric scooter users have already been observed breaking the law by riding on sidewalks and footpaths, which could increase the rate of injuries to both pedestrians and MMV riders. Furthermore, a high density of pedestrians was found to significantly reduce the speed of (e-)cyclists (*58*). Nevertheless, the effect of MMV interaction on the behavior of pedestrians in social groups and the operational efficiency of shared space under low-density conditions remains unknown. Thus, it is necessary to examine these interactions in future studies.

## Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: L. Barón, E. Freitas; data collection: L. Barón, E. Freitas; analysis and interpretation of results: L. Barón, S. Faria, E. Sousa, E. Freitas; draft manuscript preparation: L. Barón, S. Faria, E. Sousa, E. Freitas. All authors reviewed the results and approved the final version of the manuscript.

## Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## ORCID iDs

Leidy Barón https://orcid.org/0000-0002-2598-7839
Emanuel Sousa https://orcid.org/0000-0001-5128-5753

## Supplemental Material

Supplemental material for this article is available online.

## References

1. Aveni, A. F. The Not-So-Lonely Crowd: Friendship Groups in Collective Behavior. *Sociometry*, Vol. 40, 1977, p. 96.
2. Coleman, J. S., and J. James. The Equilibrium Size Distribution of Freely-Forming Groups. *Sociometry*, Vol. 24, 1961, pp. 36–45.
3. Moussaïd, M., N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz. The Walking Behaviour of Pedestrian Social Groups and Its Impact on Crowd Dynamics. *PLoS One*, Vol. 5, 2010, p. e10047.
4. Yucel, Z., F. Zanlungo, C. Feliciani, A. Gregorj, and T. Kanda. Identification of Social Relation within Pedestrian Dyads. *PLOS One*, Vol. 14, 2019, p. e0223656.

5. Zhou, L., Y. Zhao, D. Yang, and J. Liu. GCHGAT: Pedestrian Trajectory Prediction Using Group Constrained Hierarchical Graph Attention Networks. *Applied Intelligence*, Vol. 52, 2022, pp. 11434–11447.

6. Singh, H., R. Arter, L. Dodd, P. Langston, E. Lester, and J. Drury. Modelling Subgroup Behaviour in Crowd Dynamics DEM Simulation. *Applied Mathematical Modelling*, Vol. 33, 2009, pp. 4408–4423.

7. Costa, M. Interpersonal Distances in Group Walking. *Journal of Nonverbal Behavior*, Vol. 34, 2010, pp. 15–26.

8. Rajat, R., T. Ilango, and C. Satish. Design Implications of Walking Speed for Pedestrian Facilities. *Journal of Transportation Engineering*, Vol. 137, 2011, pp. 687–696.

9. Do, T., M. Haghani, and M. Sarvi. Group and Single Pedestrian Behavior in Crowd Dynamics. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2540: 13–19.

10. Rojas, F., F. Tarnogol, and H. S. Yang. Dynamic Social Formations of Pedestrian Groups Navigating and Using Public Transportation in a Virtual City. *The Visual Computer*, Vol. 32, 2016, pp. 335–345.

11. Zanlungo, F., T. Ikeda, and T. Kanda. Potential for the Dynamics of Pedestrians in a Socially Interacting Group. *Physical Review E*, Vol. 89, 2014, p. 012811.

12. Zanlungo, F., Z. Yücel, and T. Kanda. Intrinsic Group Behaviour II: On the Dependence of Triad Spatial Dynamics on Social and Personal Features; and on the Effect of Social Interaction on Small Group Dynamics. *PLoS One*, Vol. 14, 2019, p. e0225704.

13. Lu, L., C. -Y. Chan, J. Wang, and W. Wang. A Study of Pedestrian Group Behaviors in Crowd Evacuation Based on an Extended Floor Field Cellular Automaton Model. *Transportation Research Part C: Emerging Technologies*, Vol. 81, 2017, pp. 317–329.

14. Rutten, P., M. H. Lees, S. Klous, and P. M. A. Sloot. State-Space Models Reveal Bursty Movement Behaviour of Dance Event Visitors. *EPJ Data Science*, Vol. 10, 2021, p. 35.

15. D'Orazio, M., L. Spalazzi, E. Quagliarini, and G. Bernardini. Agent-Based Model for Earthquake Pedestrians' Evacuation in Urban Outdoor Scenarios: Behavioural Patterns Definition and Evacuation Paths Choice. *Safety Science*, Vol. 62, 2014, pp. 450–465.

16. Xiao, Y., J. Xu, M. Chraibi, J. Zhang, and C. Gou. A Generalized Trajectories-Based Evaluation Approach for Pedestrian Evacuation Models. *Safety Science*, Vol. 147, 2022, p. 105574.

17. Gorrini, A., S. Bandini, and M. Sarvi. Group Dynamics in Pedestrian Crowds. *Transportation Research Record: Journal of the Transportation Research Board*, 2014. 2421: 51–56.

18. Fu, T., W. Hu, L. Miranda-Moreno, and N. Saunier. Investigating Secondary Pedestrian–vehicle Interactions at Non-Signalized Intersections Using Vision-Based Trajectory Data. *Transportation Research Part C: Emerging Technologies*, Vol. 105, 2019, pp. 222–240.

19. Gorrini, A., L. Crociani, G. Vizzari, and S. Bandini. Observation Results on Pedestrian–vehicle Interactions at Non-Signalized Intersections towards Simulation. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 59, 2018, pp. 269–285.

20. Fu, T., L. Miranda-Moreno, and N. Saunier. A Novel Framework to Evaluate Pedestrian Safety at Non-Signalized Locations. *Accident Analysis & Prevention*, Vol. 111, 2018, pp. 23–33.

21. Soares, F., E. Silva, F. Pereira, C. Silva, E. Sousa, and E. Freitas. To Cross or Not to Cross: Impact of Visual and Auditory Cues on Pedestrians' Crossing Decision-Making. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 82, 2021, pp. 202–220.

22. Granié, M. A., T. Brenac, M. C. Montel, M. Millot, and C. Coquelet. Influence of Built Environment on Pedestrian's Crossing Decision. *Accident Analysis and Prevention*, Vol. 67, 2014, pp. 75–85.

23. Feng, Y., D. Duives, W. Daamen, and S. Hoogendoorn. Data Collection Methods for Studying Pedestrian Behaviour: A Systematic Review. *Building and Environment*, Vol. 187, 2021, p. 107329.

24. Shi, X., Z. Ye, N. Shiwakoti, and O. Grembek. A State-of-the-Art Review on Empirical Data Collection for External Governed Pedestrians Complex Movement. *Journal of Advanced Transportation*, Vol. 2018, 2018, p. 1063043.

25. Bohannon, R. W. Comfortable and Maximum Walking Speed of Adults Aged 20—79 Years: Reference Values and Determinants. *Age and Ageing*, Vol. 26, 1997, pp. 15–19.

26. Teknomo, K. Application of Microscopic Pedestrian Simulation Model. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 9, 2006, pp. 15–27.

27. Willis, A., N. Gjersoe, C. Havard, J. Kerridge, and R. Kukla. Human Movement Behaviour in Urban Spaces: Implications for the Design and Modelling of Effective Pedestrian Environments. *Environment and Planning B: Planning and Design*, Vol. 31, 2004, pp. 805–828.

28. Liu, Y. -C., and Y. -C. Tung. Risk Analysis of Pedestrians' Road-Crossing Decisions: Effects of Age, Time Gap, Time of Day, and Vehicle Speed. *Safety Science*, Vol. 63, 2014, pp. 77–82.

29. Rosenbloom, T. Crossing at a Red Light: Behaviour of Individuals and Groups. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 12, 2009, pp. 389–394.

30. Federici, M. L., A. Gorrini, L. Manenti, and G. Vizzari. An Innovative Scenario for Pedestrian Data Collection: The Observation of an Admission Test at the University of Milano-Bicocca. In *Pedestrian and Evacuation Dynamics 2012* (U. Weidmann, U. Kirsch, and M. Schreckenberg, eds.), Springer, Cham, 2014, pp. 143–150.

31. Hu, Y., J. Zhang, W. Song, and N. W. F. Bode. Social Groups Barely Change the Speed-Density Relationship in Unidirectional Pedestrian Flow, but Affect Operational Behaviours. *Safety Science*, Vol. 139, 2021, p. 105259.

32. Nessler, J. A., and S. J. Gilliland. Interpersonal Synchronization during Side by Side Treadmill Walking Is Influenced by Leg Length Differential and Altered Sensory Feedback. *Human Movement Science*, Vol. 28, 2009, pp. 772–785.

33. Kircher, K., O. Eriksson, Å. Forsman, A. Vadeby, and C. Ahlstrom. Design and Analysis of Semi-Controlled Studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 46, 2017, pp. 404–412.

34. Aghabayk, K., J. Esmailpour, A. Jafari, and N. Shiwakoti. Observational-Based Study to Explore Pedestrian Crossing Behaviors at Signalized and Unsignalized Crosswalks. *Accident Analysis & Prevention*, Vol. 151, 2021, p. 105990.

35. Zhang, Y., and J. D. Fricker. Investigating Temporal Variations in Pedestrian Crossing Behavior at Semi-Controlled Crosswalks: A Bayesian Multilevel Modeling Approach. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 76, 2021, pp. 92–108.

36. McNeish, D. On Using Bayesian Methods to Address Small Sample Problems. *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 23, 2016, pp. 750–773.

37. Smid, S. C., D. McNeish, M. Miočević, and R. van de Schoot. Bayesian Versus Frequentist Estimation for Structural Equation Models in Small Sample Contexts: A Systematic Review. *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 27, 2020, pp. 131–161.

38. Transportation Research Board. *Highway Capacity Manual*. The National Academies Press, Washington, D.C., 2000.

39. Johnsson, C., H. Norén, and A. Laureshyn. *T-Analyst—Semi-Automated Tool for Traffic Conflict Analysis*. InDeV, Horizon Project, 2020.

40. Pokorny, P., B. Skender, T. Bjørnskau, and M. P. Hagenzieker. Video Observation of Encounters between the Automated Shuttles and Other Traffic Participants along an Approach to Right-Hand Priority T-Intersection. *European Transport Research Review*, Vol. 13, 2021, p. 59.

41. Bulla-Cruz, L. A., A. Laureshyn, and L. Lyons. Event-Based Road Safety Assessment: A Novel Approach Towards Risk Microsimulation in Roundabouts. *Measurement*, Vol. 165, 2020, p. 108192.

42. Johnsson, C., A. Laureshyn, and C. Dágostino. A Relative Approach to the Validation of Surrogate Measures of Safety. *Accident Analysis & Prevention*, Vol. 161, 2021, p. 106350.

43. Tsai, R. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using off-the-Shelf TV Cameras and Lenses. *IEEE Journal on Robotics and Automation*, Vol. 3, 1987, pp. 323–344.

44. Pinheiro, J. C., and D. M. Bates. Linear Mixed-Effects Models: Basic Concepts and Examples. In W. K. Härdle (ed), *Mixed-Effects Models in S and S-PLUS. Statistics and Computing*. Springer, New York, NY, 2000, pp. 3–56.

45. Singh, S. G., and S. V. Kumar. Dealing with Multicollinearity Problem in Analysis of Side Friction Characteristics Under Urban Heterogeneous Traffic Conditions. *Arabian Journal for Science and Engineering*, Vol. 46, 2021, pp. 10739–10755.

46. Monsalves, M. J., A. S. Bangdiwala, A. Thabane, and S. I. Bangdiwala. LEVEL (Logical Explanations & Visualizations of Estimates in Linear Mixed Models): Recommendations for Reporting Multilevel Data and Analyses. *BMC Medical Research Methodology*, Vol. 20, 2020, pp. 1–9.

47. R Core Team. R: A Language and Environment for Statistical Computing. https://www.r-project.org/.

48. Montufar, J., J. Arango, M. Porter, and S. Nakagawa. Pedestrians' Normal Walking Speed and Speed When Crossing a Street. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2002, 2007, pp. 90–97.

49. Uzzell, D., and N. Horne. The Influence of Biological Sex, Sexuality and Gender Role on Interpersonal Distance. *The British Journal of Social Psychology*, Vol. 45, 2006, pp. 579–597.

50. Fu, L., S. Cao, Y. Shi, S. Chen, P. Yang, and J. Fang. Walking Behavior of Pedestrian Social Groups on Stairs: A Field Study. *Safety Science*, Vol. 117, 2019, pp. 447–457.

51. Ishaque, M. M., and R. B. Noland. Behavioural Issues in Pedestrian Speed Choice and Street Crossing Behaviour: A Review. *Transport Reviews*, Vol. 28, 2008, pp. 61–85.

52. Holland, C., and R. Hill. The Effect of Age, Gender and Driver Status on Pedestrians' Intentions to Cross the Road in Risky Situations. *Accident Analysis & Prevention*, Vol. 39, 2007, pp. 224–237.

53. Díaz, E. Theory of Planned Behavior and Pedestrians' Intentions to Violate Traffic Regulations. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 5, 2002, pp. 169–175.

54. Fitzpatrick, K., M. A. Brewer, and S. Turner. Another Look at Pedestrian Walking Speed. *Transportation Research Record: Journal of the Transportation Research Board*, 2006. 1982: 21–29.

55. Sorokowska, A., P. Sorokowski, P. Hilpert, K. Cantarero, T. Frackowiak, K. Ahmadi, A. M. Alghraibeh, et al. Preferred Interpersonal Distances: A Global Comparison. *Journal of Cross-Cultural Psychology*, Vol. 48, 2017, pp. 577–592.

56. Kadali, B. R., and P. Vedagiri. Review of Pedestrian Level of Service: Perspective in Developing Countries. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2581: 37–47.

57. El-Agroudy, M., H. Abou-Senna, and E. Radwan. Mobility-as-a-Service: Simulation of Multi-Modal Operations in Low-Density Cities. *Transportation Research Record: Journal of the Transportation Research Board*, 2022. 2676: 235–246.

58. Zhang, C., B. Du, Z. Zheng, and J. Shen. Space Sharing between Pedestrians and Micro-Mobility Vehicles: A Systematic Review. *Transportation Research Part D: Transport and Environment*, Vol. 116, 2023, p. 103629.