ORIGINAL PAPER

Pedestrian gaze pattern before crossing road in a naturalistic traffic setting

Qingxian Zhao[1], Xiangling Zhuang[1*], Tong Zhang[1], Yunqi He[1] and Guojie Ma[1]

Abstract

Background Gaze is the primary way for pedestrians to obtain clues from traffic scenes before making decisions. Therefore, understanding pedestrian gaze pattern is vital for traffic safety in general and for the design of autonomous vehicles.

Methods In this study, participants made road-crossing decisions in a naturalistic traffic scene, with an eye-tracker recording their gaze behaviors. We manually encoded the recorded videos with 14,898 fixations, and then analyzed the gaze pattern at three levels from general to specific: gaze towards overall scenes, gaze towards vehicles and gaze towards components of vehicles.

Findings At the first level, our findings indicate that frequent fixations began to appear at the distance of 100 m and peaked around 5–30 m away from pedestrians. Transversely pedestrians mainly gazed at the two lanes adjacent to themselves. Pedestrians allocated 53% gaze duration to motor vehicles. For a specific vehicle, which is the second level, the gaze duration varied with vehicles' attributes such as distances, sizes, and types. Finally, at the third level, we discovered that pedestrians' gaze duration on different vehicle components varied with the longitudinal distance. As vehicles approach, the main area of fixation expanded from the near side headlight to the whole front and near side, and finally shift to the near side of a vehicle.

Implications The distribution of fixations in space and vehicle components before pedestrian crossing can provide fundamental information for understanding and modeling of pedestrian's road-crossing behaviors. In practice, our findings can guide the timing and position of information displays on autonomous vehicles to facilitate friendly interaction with pedestrians.

Keywords Traffic safety, Human-vehicle interaction, Eye-tracker, Pedestrian road-crossing decision, Autonomous vehicles

1 Introduction

Pedestrian safety has been challenged worldwide with the rapid development of motorization. In 2018, about 310,000 pedestrians were killed on road worldwide, accounting for 23% of all deaths occurred in traffic accidents [1]. To protect pedestrians, researchers have made various efforts to understand pedestrian behaviors as well as their decision-making process before accidents occurred. This study focuses on pedestrian behavior while crossing a road because compared with other pedestrian tasks (e.g. finding route, avoiding obstacles), it often involves interaction with vehicles, thus exposing higher risks to pedestrians [2].

From a pedestrian's perspective, crossing a road involves complex cognitive processes. Some researchers proposed to divide the road-crossing task into *pre-crossing* and *crossing* [3]. Instead of directly stepping into the street, pedestrians conduct multiple mental processes in the *pre-crossing* stage, such as observation, perception,

*Correspondence:
Xiangling Zhuang
zhuangxl@snnu.edu.cn
[1] Shaanxi Key Laboratory of Behavior and Cognitive Neuroscience, School of Psychology, Shaanxi Normal University, Xi'an, China

Zhao *et al. European Transport Research Review*   (2023) 15:31

Page 2 of 16

judgement and decision-making, to determine where and when they should cross. The process is also captured by the model of situation awareness [4], where pedestrians' crossing decisions and behavior depend on situation awareness. They *perceive* the state of vehicles and other road users, *comprehend* the current situation, and *predict* future states to make decisions. The primary explicit behavior of these processes is manifested by pedestrians' gaze behavior, which is the entry point for us to understand the characteristics of pedestrians' pre-crossing stage and scene perception.

In other tasks beyond crossing the road, gaze behavior has been shown to provide rich cues on individuals' intentions and help understand the higher-level events [5]. In this study, we concretized pedestrians' gaze behavior as the gaze pattern by operationalizing it as *the fixation characteristics towards road elements under various contexts*. More specifically, we focus on the targets and duration of pedestrian's fixations while making a road-crossing decision. Understanding gaze pattern is important for both pedestrians' safety and vehicle design. First, a complete profile of the gaze pattern can facilitate future modeling of pedestrian decision-making and behavior, especially on what information pedestrians rely on while making decisions and how they interact with vehicles (e.g., the range and type of vehicles). Second, it has been a common practice to equip an external human–machine interfaces (eHMI) on autonomous vehicles to facilitate communication between vehicles and pedestrians. However, the main user of eHMIs is pedestrian, so the timing and position of information presentation on eHMIs should consider pedestrian gaze pattern. Therefore, the gaze pattern will inform the pedestrian-friendly design of autonomous vehicles [6, 7].

2 Pedestrian gaze pattern before crossing the road

Previous researchers have conducted some experiments and field studies to explore pedestrian gaze patterns directly, or more general behaviors in scene perception that might indirectly inform gaze patterns. To integrate these findings, we divided the traffic scene into three levels from general to specific: road level, vehicle level, and vehicle's components level. The following three sections overviewed previous findings in these three levels to capture pedestrians' possible gaze patterns in real and complex traffic scenes.

2.1 First level: gaze pattern for the overall scene

Pedestrians must explore the local environment to extract and interpret information from roads in several ways before crossing, especially visually. However, natural traffic situation contains complex elements and clues, while pedestrian's attention resource is limited.

Therefore, instead of scrutinizing every road element with equal attention, pedestrians will likely distribute their visual attention only to the road elements critical to their tasks.

The inclination of selectively gazing at relevant elements has been seen in other contexts. For example, when participants walked to complete different tasks, they made a saccadic eye movement to align with the clues related to their following action. For example, participants' fixations mainly fell on their next point of footfall [8, 9], or on the regions and objects which have the most information about their task [10].

In the traffic scenario, gaze behavior is also goal-directed. Some experiments in the traffic scenes have proved that fixations are directed to some task-relevant objects and areas while the irrelevant others on the road are ignored [11]. For example, pedestrians paid most attention to the path when they walked along a pre-defined route [12]. At the signalized plus intersection, incompliant pedestrians mainly fixated on cars while the compliant pedestrians who waited for the green light mainly fixated on traffic lights [13]. These pieces of evidence demonstrate that pedestrians do have gaze patterns compatible with their tasks and contexts. The variability of gaze patterns also calls for the exploration of gaze patterns in various traffic scenes to understand pedestrian decision-making systematically.

In this study we choose to explore pedestrian gaze patterns at an uncontrolled multi-lane road, which is still missing in literature. Compared with the controlled scenarios at the signalized crosswalk [13], the uncontrolled multi-lane road has much more complex information to process. The decision-making is purely based on pedestrian situation awareness with no help from traffic rules. Besides, prior studies mainly focused on the type of the fixation target, while the spatial distributions of fixations are also indispensable aspects of gaze patterns. To fill the gap, we will record both the target and the position of pedestrians' fixations when they deciding whether to cross a multi-lane road without traffic controls.

2.2 Second level: gaze pattern towards vehicles

As stated, pedestrians intentionally select task-relevant objects and areas from the overall traffic scene. Since vehicles are the primary sources of risk in road crossing tasks, they become the essential targets of fixations relative to other road elements. However, similar to the selection at the road level, pedestrians may not distribute equal attention to all vehicles in their visual field. Instead, their fixations towards vehicles can be determined by specific attributes of vehicles such as distance and type.

First, the position of a vehicle is one of the main factors that determine whether it will attract visual attention.

The characteristics of the human visual system limit its observable distance range. Even for the vehicles in the visible field of view, distance also affects the probability of being looked at. A prior study found that cars at a distance were fixated more than cars already in the crosswalk (i.e., near to pedestrians), both at intersections and roundabouts [13]. In a simulated setting with only one vehicle, pedestrians' fixations on the vehicle gradually increased as it approached from 30 m [14]. These findings reveal that vehicle's distance is a vital factor affecting pedestrians' gaze patterns towards vehicles. However, since fixation results from selective attention, the distance range of pedestrian's fixations may likely be different in more complex road conditions where pedestrians have more vehicles to select. Natural traffic scenarios usually have multiple vehicles and multiple lanes. Correspondingly, our study aims to determine the spatial distribution of fixations towards vehicles in both longitudinal and traverse directions.

Second, some external features of the vehicle may also determine how likely the vehicle is to be looked at. Larger vehicles are easier to gaze than smaller ones because they require less effort to keep in the fovea [15]. While exploring how vehicle size affected speed perception, Clark et al. [16] found pedestrians' eye movements towards vehicles changed with vehicle size because pedestrians' fixations are mainly concentrated around the visual centroid. Although no evidence is available on how color affects pedestrian gaze behavior, color affects physiological and psychological processes in drivers [17]. Besides, some colors, such as black, blue, grey, green, red, and silver, are associated with higher crash risk than white because of poor visibility [18]. By analogy, we expect vehicles' colors may also affect pedestrians' gaze behavior when they try to get clues from vehicles. Therefore, we will analyze how pedestrian gaze pattern differs across different vehicle positions, types and colors while making a road crossing decision.

2.3 Third level: gaze pattern towards vehicle's components

We have introduced the pedestrian's gaze pattern towards the overall scene and vehicles, and then we continue to narrow the scope further to the vehicle's components. In other words, how does pedestrian gaze behavior change across different components of a vehicle? The answer offers clues to how pedestrians gather cues from vehicles for decision-making. For example, a pedestrian constantly gazing at the front window may be seeking to form eye contact with a driver.

More importantly, fixation patterns across vehicle components can inform the design of autonomous vehicles (AVs). With the absence of human drivers in AVs, researchers proposed to equip external human–machine interfaces (eHMIs) to display AVs' intention to increase the efficiency of interactions [6]. The question followed is where the eHMIs should be placed. The study on the position of eHMIs is essential because the display location may affect pedestrians' crossing intentions and behaviors [19]. One approach to select the position of the display is to evaluate them at different positions. Several positions for eHMIs have been proposed, including the window screen [20], the front of cars [6] and so on. They were also evaluated by questionnaire [6] or focus group discussion [20].

While these approaches are straightforward, they may result in an overly rigid and straightforward choice of display position. Pedestrians may look at different positions in different conditions. For example, pedestrians walking in a parking lot often looked at the back of parked cars for the brake lights and turning lights to predict the movement of cars [21]. Similarly, one study found pedestrians fixated mainly on the car's bumper in the distance of 25–30 m, while the fixation on the hood remained high in the distance of 20–30 m. As the distance decreased to 5–20 m, pedestrians' gaze patterns shifted significantly to the windshield [14]. However, in this study [14], the only car presented was driven by the experimenter at 50 km/h on a straight one-lane road. Participants were required to look at the car and indicate their willingness to cross. In this task, pedestrians continuously looked at the only car without attention shift. However, in the complex traffic scenes, pedestrians may shift their attention to road elements, other vehicles, and even scenes unrelated to road-crossing tasks. For example, pedestrians would change their subsequent behavior based on the number of approaching vehicles [22], thus, we expect their task-relevant gaze behaviors to be affected by other vehicles and even other road elements. In this study, we aim to determine pedestrians' primary areas of interest towards vehicles' components in a traffic scene containing as many complex elements as possible. We assume that the components where pedestrians spontaneously prefer to look at under natural scenes may be a better location for eHMIs, because they are more in line with pedestrians' needs and expectations.

2.4 The aim of this study

To sum up, identifying pedestrian gaze behavior towards vehicles is essential both for theory and practice. Previous studies have described pedestrians' gaze behavior under different tasks and scenes. However, a complete portrait of pedestrian gaze patterns still entails systematic measurements of pedestrians' gaze patterns at the three levels we identified in natural settings. Therefore, the main objective of this study is to explore pedestrian

Zhao *et al. European Transport Research Review*   (2023) 15:31

Page 4 of 16

gaze patterns before crossing a multi-lane road in naturalistic traffic scene.

Specifically, we recorded pedestrian gaze behavior to answer the following questions at three levels (Fig. 1). Firstly, on the overall scene level, we aimed to identify pedestrians' gaze targets and describe the overall spatial distribution of fixations. Secondly, on the vehicle's level, we went further understand how certain vehicles' attributes such as position, size, and color may affect pedestrian gaze behavior. Finally, we tried to portray the dynamic gaze pattern of pedestrians towards vehicles' components in different conditions.

3 Methods

3.1 Participants

Seventeen participants (nine females, eight males) recruited from the Shaanxi Normal University took part in this study. Although the sample size is not large, it is common in studies that involve eye-tracking of dynamic scenes [13], N=12, [10], N=7) because the data analysis requires time-consuming manual coding of eye-tracking records. The participants were aged between 20 and 24 years (M=21.53, SD=1.14). Since road crossing is a basic daily activity that does not require special training, we believe this age group can reflect the general gaze pattern across different age groups except for some children and elderly pedestrians. All participants reported normal or corrected-to-normal vision. In addition, they signed informed consent after a detailed description of the study and received monetary compensation for the participation.

3.2 Apparatus

To simultaneously record pedestrians' gaze behaviors and the scene, eye movements were recorded at 100 Hz using the Tobii Pro Glasses 2 mobile eye-tracker. The eye-tracker uses infrared light to record human pupil and corneal reflexes. As a result, direct sunlight may affect the accuracy of recording. To avoid the disturbance from bright sunlight, we conducted our study in the early morning and late afternoon on 13 days. Each time, we performed the 1-point calibration procedure to calibrate the eye tracker. The Tobbi Pro Lab was used for calibration and replay. We used the Attention Filter in Tobii Pro Lab, which is the Tobii Pro IV-T Filter, with the velocity threshold parameter set to 100°.

3.3 Site of study

Our study aims to record pedestrian natural gaze pattern before crossing at an uncontrolled site. The observation was conducted on Zhuque street, a typical six-lane two-way road in Xi'an, China (Fig. 2). The participants' imaginary destination is the yellow line in the middle of the road that separated traffic into two ways. To get to the destination, the participants must cross three lanes, where vehicles were coming from the participants' left. At the selected site is 16 m away from the nearest zebra crossings, so the vehicles were moving in a steady speed
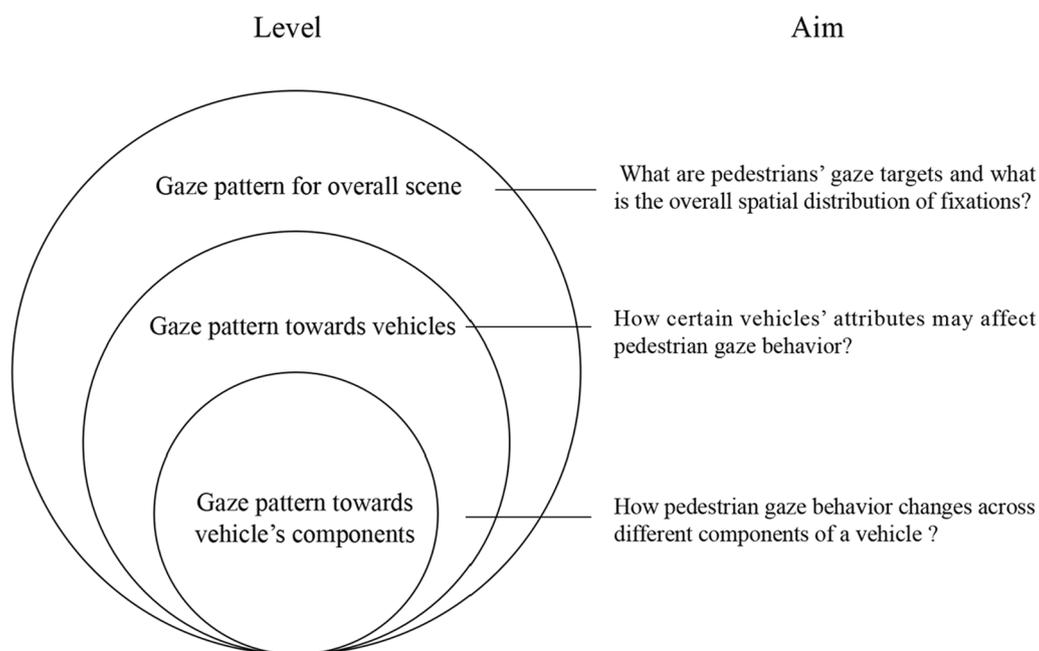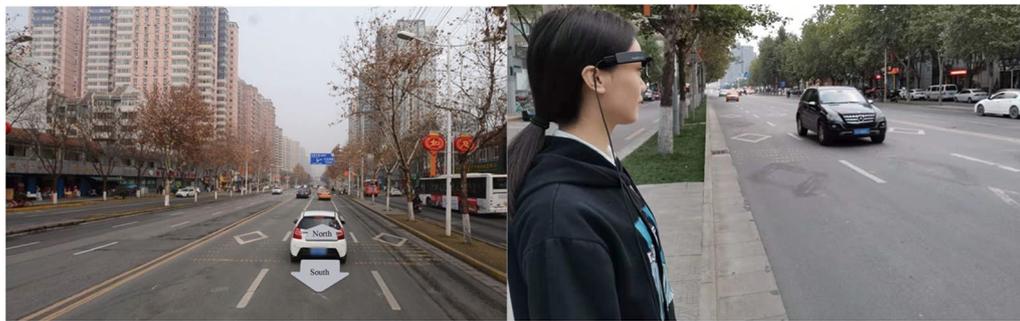

Fig. 1 The aims of this study

Zhao *et al. European Transport Research Review*   (2023) 15:31

Page 5 of 16

a. The site of this study   b. A participant wearing the eye-tracker

Fig. 2 Illustration of the experimental setting

around 30–50 km/h, which is the typical speed on urban road. The traffic flow on the three lanes of the road is 1280 vehicles each hour. It is a normal traffic flow of a six-lane urban arterial road.

3.4 Procedure

On arrival, participants were briefed with the experimental instructions. After fitting and calibrating the eye tracker, participants were led to the edge of the pavement where the experiment was conducted (Fig. 2). Participants' task was to make road-crossing decisions as in their daily life. Except that the destination is set at the yellow line in the middle of the road. We proceeded each trial in three steps to simulate daily road crossing decision-making and record the natural gaze behaviors when participants interacted with vehicles.

3.4.1 Step 1: Wait

Participants stood at the curb with their back to the street. Herein they could only hear the traffic noise but could not see any vehicles. Since vehicles are the most important objects of pedestrian observation, participants will quickly decide to cross if there is no vehicle on the road, resulting in a short decision phase and a small number of recorded fixations. Therefore, the experimenter observed the traffic conditions and issued the start order only when there were vehicles on the road, but the number and distance of vehicles in all trials are totally random.

3.4.2 Step 2: Observe

When the participants heard the "action" command from the experimenter, they turned around and faced the road. To cross the road safely, participants observed the traffic condition until they identified the appropriate chance to cross. Fixation data collected from this step was subjected to subsequent analysis.

3.4.3 Step 3: Decide

The third step is to "decide" to cross. In daily life, the start of walking marks the completion of the decision-making. To ensure safety, all participants were asked to turn around instead of walking when they thought it was time to cross. The procedure is common in studies focusing on pedestrian crossing decision-making safety (e.g. [14]).

In short, one trial of decision-making was composed of "back to the road—turn around after command—turn around when safe." Before the formal trials, participants performed training trials until they fully understood the process. They could stop at any time when they want to take a break. Overall, each participant made 30 decisions.

3.5 Data extraction and analysis methods

In this study, the eye-tracker recorded videos with fixation points. By replaying these videos, we can locate any frame from the scene and corresponding fixation information. To describe the characteristics of participants' gaze patterns along the waiting process, we manually extracted the fixation information based on a predefined list of variables (Table 1). Before coding, we discussed with experts to clarify the coding standards of these variables. All the fixations were coded by one experimenter.

For each fixation, we coded the target and duration of the fixation. More importantly, we need to extract the target's position (longitudinal and transverse distance) and its color, size, type, and specific components if the target is a vehicle. To ensure the validity in the coding of vehicle distance, we measured the distance of all readily discernible markings on the road, such as trees on the side of the road, traffic signs on the ground, etc. Using a consistent estimation criterion of one experimenter, we can roughly calculate the range of gaze point distances relative to these markings.

To extract which vehicle components a fixation is located, every vehicle was manually divided into twelve

Table 1 Coding information description for each fixation

| Coding categories	Explanation
Trial ID	Identify which decision-making trial the fixation belongs to;
Target ID	Identify the gazed target;
Target type	Record the type of the gazed object; motor vehicle = 0; Non-motor vehicle = 1; Other road elements = 2
Dist_Longitudinal	The actual longitudinal distance of the fixation
Dist_transverse	The actual transverse distance of fixation
Gaze duration	Duration of the fixation recorded by eye-tracker
Vehicle size	Normal = 0 (like sedan and taxi); big = 1 (like bus and truck)
Vehicle type	Divided by vehicle function; Car = 0; bus = 1; taxi = 2
Vehicle color	White and silver = 0; black = 1; green = 2; yellow = 3; red = 4; blue = 5
Vehicle components	We divided vehicles into 12 different components. Figure 3 shows these components
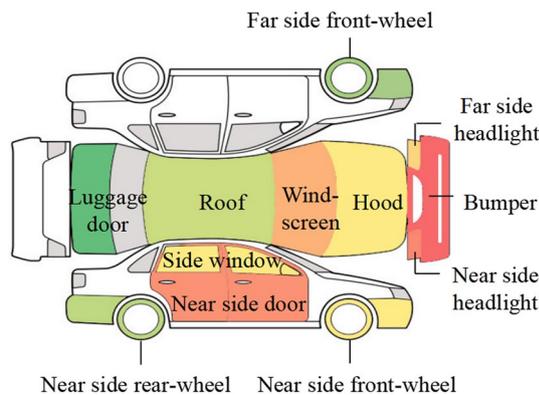
Fig. 3 The division of vehicle components

areas of interest according to different components (Fig. 3).

After coding, we got a dataset composed of 14,989 fixations. Since fixations may be affected by participant-specific variables, data from different measurements are inter-dependent rather than independent. To address the heterogeneity between participants and the correlation between multiple data from the same participant, we choose to build linear *mixed* models (LMM). The principle of the linear mixed model is as follow: $\mathbf{Y} = \gamma\, \mathbf{u} + \delta\, \mathbf{v} + \varepsilon$, where $\mathbf{u}$ is the fixed effect, $\mathbf{v}$ is the random effect, γ and δ are parameters and $\delta \sim (0, \sigma^2)$, ε is a residual error. Different with linear models, a LMM can estimate a model for each participant by setting a different intercept or slope for each participant. In this study, we set a random intercept for each participant by putting "participant" into the random effect. It allowed us to examine the effect of other independent variables by assuming a different "baseline" for each participant.

We used R (R Core Team, 2016) and its lme4 package [23] to estimate all the models.

4 Results

We collected 14,989 fixations that lasted 4876.76 s, while the total observing duration is 9417.18 s. The sample rate of our total data ranges from 65 to 96% (see Additional file 1 for more details), which is the acceptable range for outdoor eye-tracking experiments.

The aim of this study is to concretize pedestrians' gaze patterns with fixation characteristics, including gaze duration and the external features of gazed vehicles. We organized our results along three levels to answer the questions mentioned earlier (Fig. 1).

4.1 Gaze pattern for the overall scene

We firstly analyzed the gaze duration of 14,989 fixations (Fig. 4). Most single fixations lasted less than 2 s (99.3%) and centralized around 0.07–0.46 s (81.22%).

We calculated the total and mean gaze duration under different longitudinal and transverse distances to describe the spatial distribution of pedestrian fixations for the overall scene. The total gaze duration is the sum of the fixation duration within the same distance range. In contrast, the mean duration is calculated by dividing the total duration by the number of fixations at each distance.

The heat maps (Figs. 5 and 6) was an abstract overhead view of a real road and visualized the gaze duration at different positions to support an intuitive overview of the gaze pattern. In the heat maps, red and yellow indicate where participants spent a long time observing, while green indicates areas that received shorter fixations.

Longitudinally, we can gauge that the long total gaze duration is mainly concentrated within the range of 5–30 m. For fixations 40 m away, the number of fixations
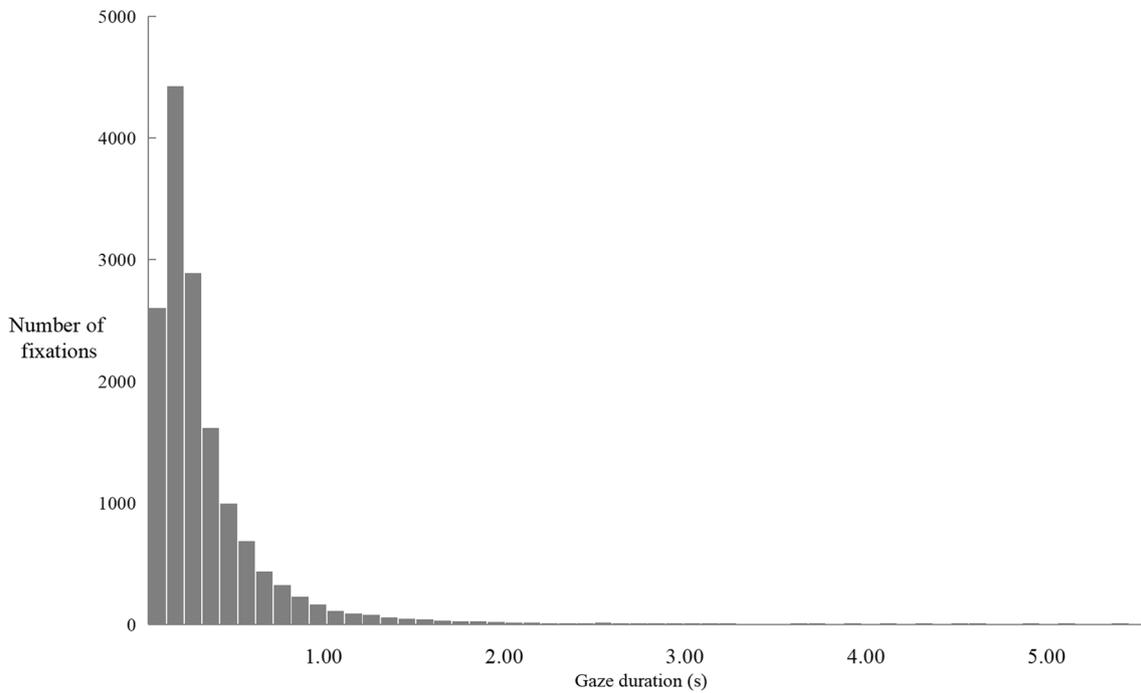
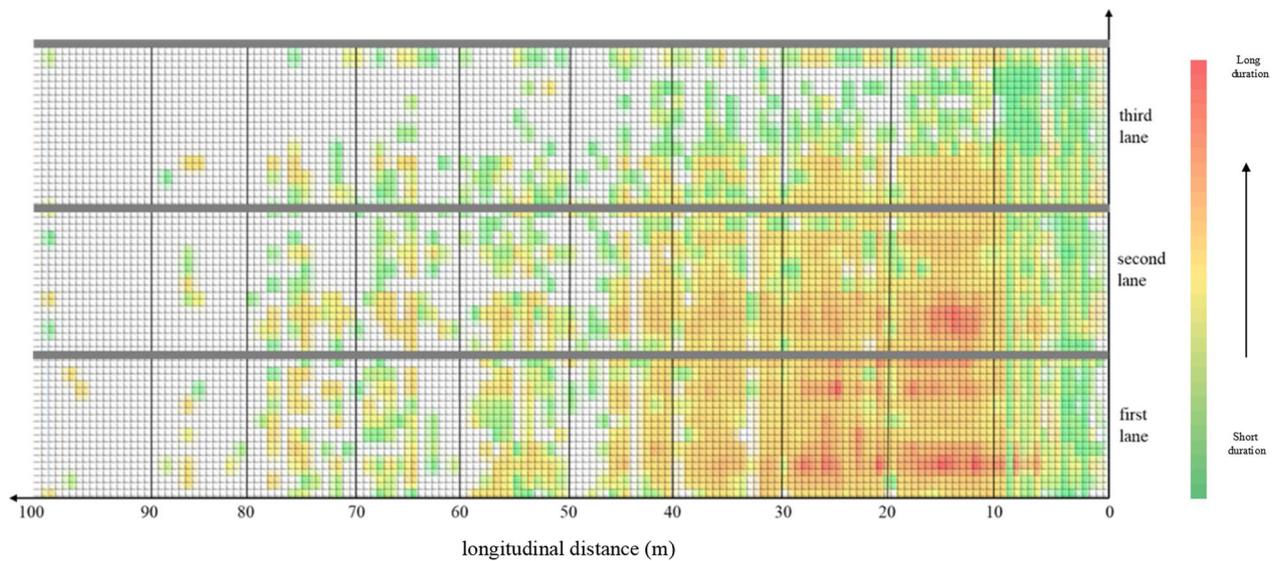**Fig. 4** The distribution of fixations' gaze duration

Fig. 5 Heat map of *total* gaze duration under different distances

become smaller, but their mean duration become longer. The "smaller" and "longer" are only qualitative statements judged by the color, because it is challenging for us to demonstrate the gaze duration quantitatively for each continuous distance. In the transverse direction, fixations are mainly concentrated in the first two lanes adjacent to pedestrians.

To confirm the intuitive estimation of trend, we plotted the average fixations duration relative to longitudinal distances. To further compare the difference in gaze duration at different distance, we re-coded the continuous longitudinal distance into the group variable by dividing 0–100 m into ten groups at intervals of five meters. The last group is 45–100 m. In the transverse
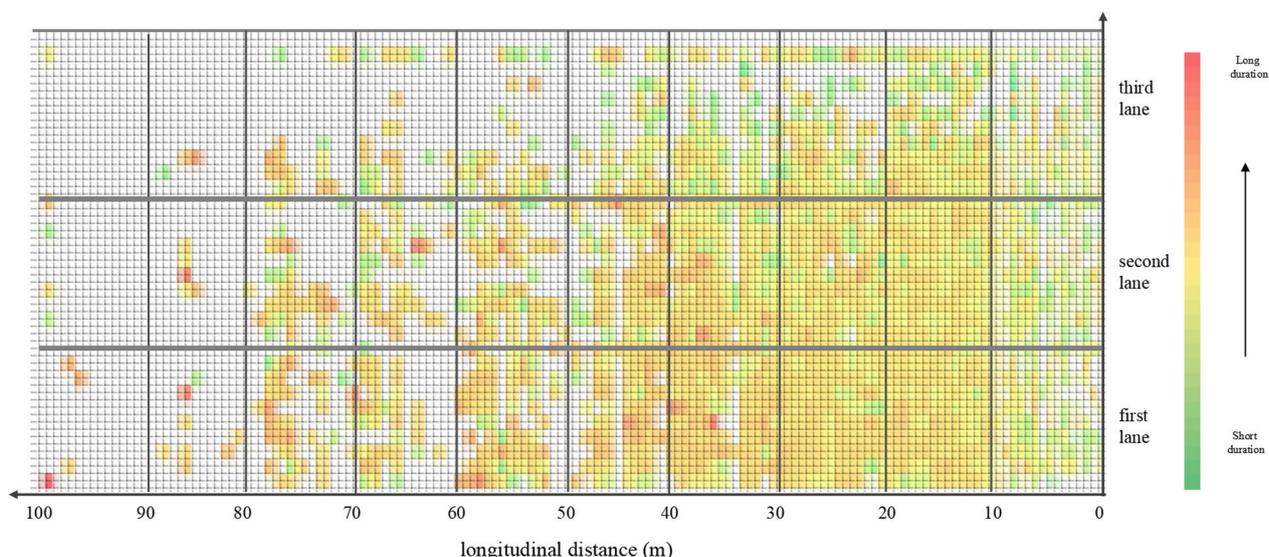
Fig. 6 Heat map of *mean* gaze duration under different distances

direction, we divided them into three lanes according to the existing road lane signs (see Figs. 7 and 8). Figure 7 shows that the total gaze duration gradually increases from 30–100 m and remains high within the range of 5–30 m before it finally decreases between 0–5 m. For example, in the first lane, the average total gaze duration in these three intervals is 158.6, 410.3, and 67.7 s. Figure 8 indicates that participants' mean gaze duration towards all targets increases with distance, which means the farther area has fewer fixations but a longer mean duration. For example, in the first lane, mean gaze duration almost monotonically decreases from

0.49 to 0.23 s within the range of 0–100 m. Overall, the changes of total and mean gaze duration in the longitudinal is more evident in near lanes than farther lanes (esp. the third lane). Integration of the two figures portrayed a gaze profile with long and stable fixations at a farther distance and frequent fixations with short duration at a near distance.

In addition to the spatial distribution of the fixations, the targets of fixations were also analyzed. Participants observe three types of targets: motor vehicles, non-motor vehicles and other road elements. Table 2 shows the percentage of gaze duration on them. Participants allocate
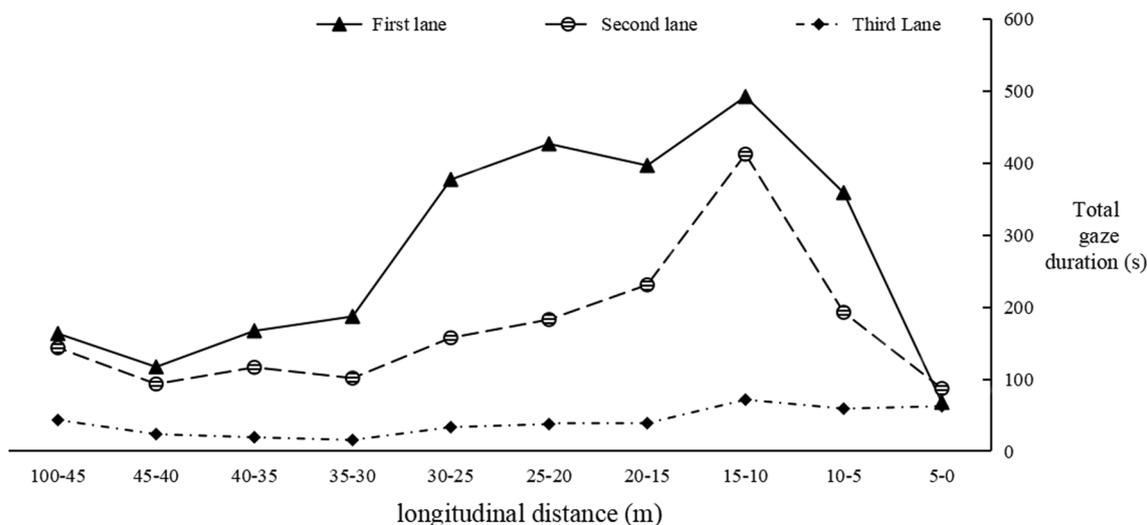

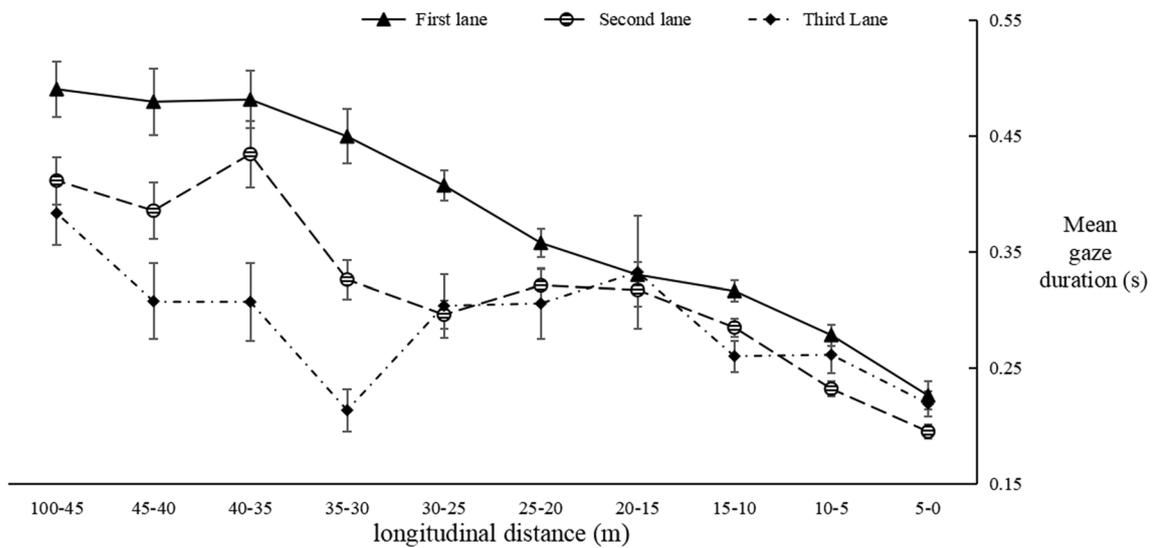**Fig. 7** Participants' *total* gaze duration as functions of distances

Fig. 8 Participants' **mean** gaze duration as functions of distances

Table 2 Distribution of gazed target towards different road elements

	Mean fixation duration (SD)	Total fixation duration (s)	Percentage of total duration (%)
Motor vehicles	0.32 (0.36)	2584.64	53.0
Non-motor vehicles	0.41 (0.49)	120.87	2.5
Other road elements	0.32 (0.35)	2171.24	44.5

53% of their total time to motor vehicles and 45% to other road elements.

4.2 Gaze pattern towards vehicles

Since vehicles are one of pedestrians' main gazed targets, we followed our plan to analyze the gaze pattern at the second level: the vehicles. To that end, we added up the gaze duration of all fixations belonging to the same vehicle. We aimed to identify how vehicle features such as position, color, size, and type affected the fixations. In total, we coded 7956 fixations fell on the vehicles, but because there were multiple consecutive fixations on the same vehicle, the number of gazed vehicles was 5750.

Table 3 The numbers of vehicles and fixations towards them[a]

Vehicle's features		Number of fixations	Number of vehicles	%	Average number of fixations	Mean fixation duration
Size	Normal	7697	5593	97.3	1.38	0.52
	Big	259	157	2.7	1.65	0.63
Type	Car	6714	4940	85.9	1.36	0.50
	Bus	35	26	0.5	1.35	0.55
	Taxi	1207	784	13.6	1.54	0.69
Color	Black	1284	959	16.7	1.34	0.47
	Green	467	318	5.5	1.47	0.65
	Yellow	822	529	9.2	1.55	0.70
	Red	229	178	3.1	1.29	0.49
	Blue	131	86	1.5	1.52	0.55
	White and silver	5023	3680	64.0	1.36	0.51

[a] The *average number of fixations* is calculated by dividing the number of fixations by the number of vehicles; the *mean fixation duration* is calculated by dividing the total gaze duration by the number of vehicles

Table 3 shows the number of fixations and gazed vehicles with different features, the average number of fixations towards them, and their mean fixation duration.

We also drew the heat maps to describe the total and mean gaze duration distribution for the fixations falling on vehicles (Figs. 9 and 10). Notice that although the two figures appear similar to Figs. 5 and 6 (include all fixations fell on the road and vehicles), they only include the fixations fell on the vehicles. Although the fixation data towards vehicles is only a subset of the fixation towards the overall scene, the distributions across different lanes and distances are similar. The figures using total (Figs. 9 and 11) and mean gaze duration (Figs. 10 and 12) as indicators still show the opposite trend. The longer total fixations are mainly concentrated within 5–20 m, while the longer mean fixations are within 25–90 m.

Figure 11 shows that the total gaze duration on vehicles increased and reached the highest point at the 10–15 m interval, then showed the downward trend within the range of 15–40 m. Finally, from 40 to 100 m, they rose slightly with small slopes, regardless of lanes. Figure 12 is similar to Fig. 8; they both showed that
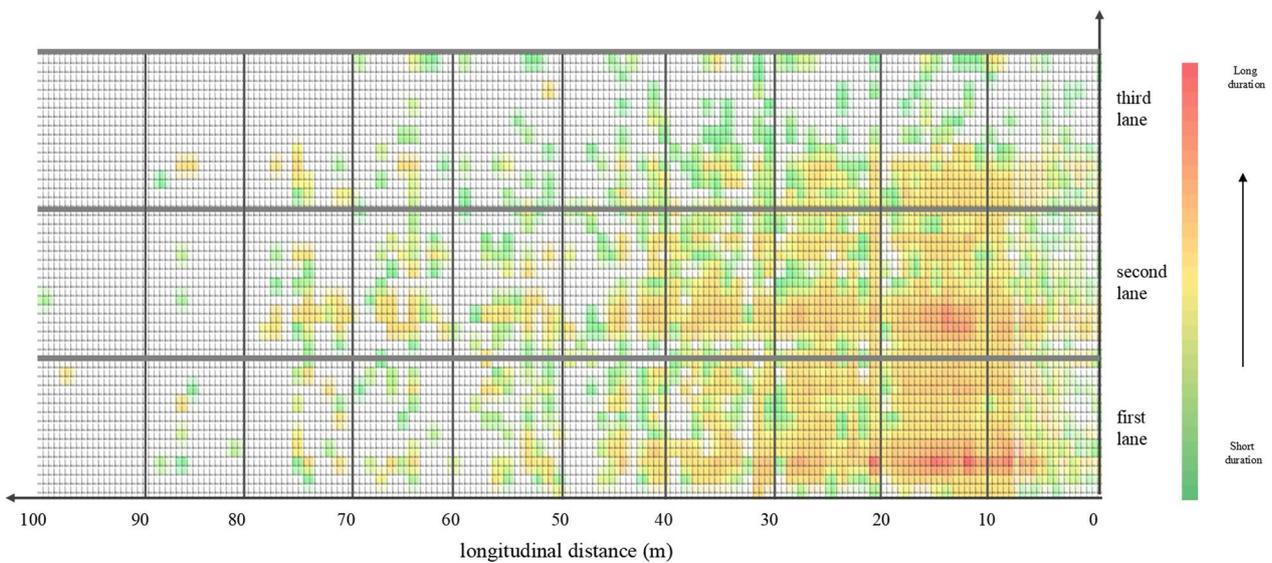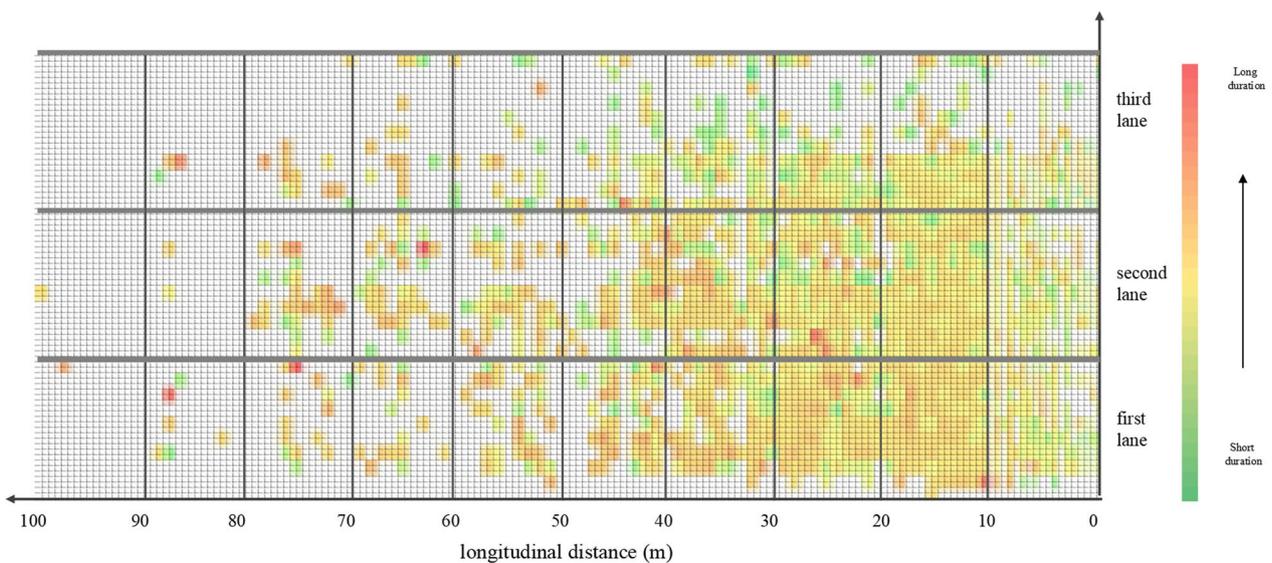

Fig. 9 Heat map of *total* gaze duration *on vehicles*
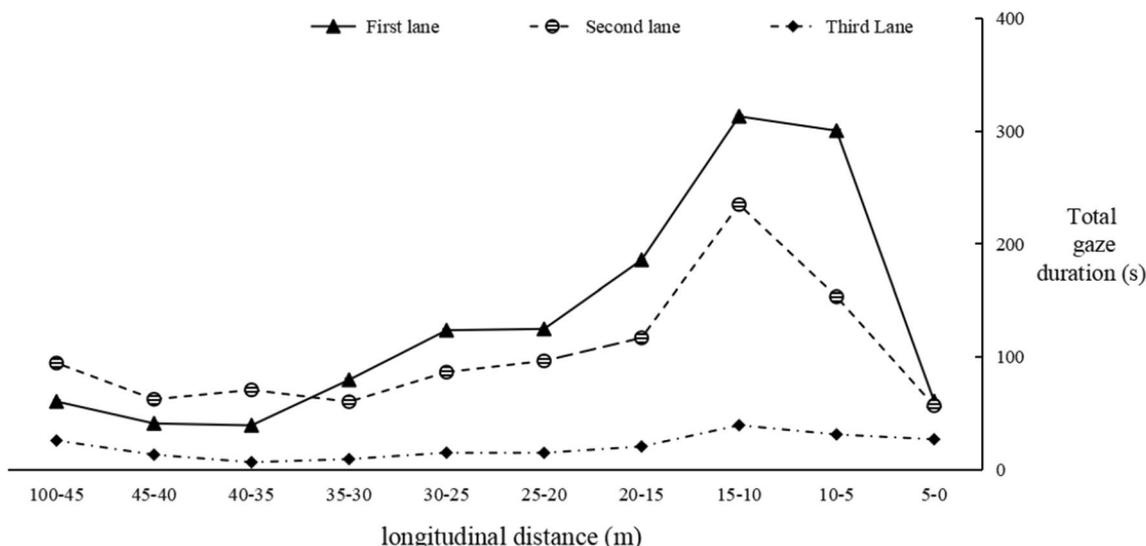

Fig. 10 Heat map of *mean* gaze duration *on vehicles*

Zhao *et al. European Transport Research Review*   (2023) 15:31

Page 11 of 16

Fig. 11 Participants' **total** gaze duration as functions of distances **on vehicles**
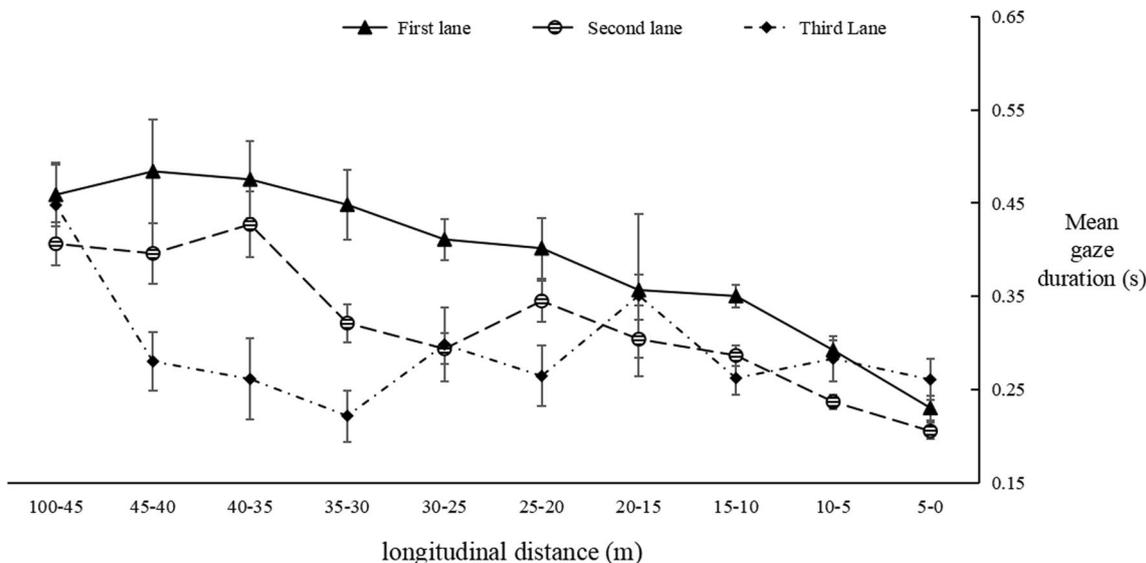

Fig. 12 Participants' **mean** gaze duration as functions of distances **on vehicles**

the mean gaze duration is longer at a farther distance, regardless of whether the gazed targets are all objects on the road (Fig. 8) or only vehicles (Fig. 12).

Besides vehicle position, other vehicle features such as size and type also affected vehicle gaze duration. To analyze how these features affected gaze duration, we built LMM models to analyze their effects. We first built Model 1 with fixed effects of size and type, with a random effect of participant. The core syntax of Model 1 is: *Total gaze duration ~ Size + Type + (1|Participant).*

Similarly, we built a Model 2 with an extra fixed effect of color. The core syntax of Model 2 is: *Total gaze duration ~ Size + Type + Color + (1|Participant).* Comparison of the two models yielded no improvement in fitness with the additional components (χ^2 (5) = 6.4178, p = 0.2677). That means color does not have a significant effect on the total gaze duration. The final fixed effect results are displayed in Table 4. The result revealed that both size and type of vehicle significantly affected gaze duration. For example, participants gave significantly longer gaze

Table 4 Fixed effects of size and type

	Estimate value (SE)	df	t value
Intercept	0.48 (0.03)	15.66	14.74[***]
Size-big (normal as baseline)	0.12 (0.05)	5734.47	2.44[*]
Type-bus (car as baseline)	− 0.03 (0.12)	5745.83	− 0.23
Type-taxi (car as baseline)	0.21 (0.02)	5741.20	9.07[***]

*$p < 0.05$, ***$p < 0.001$
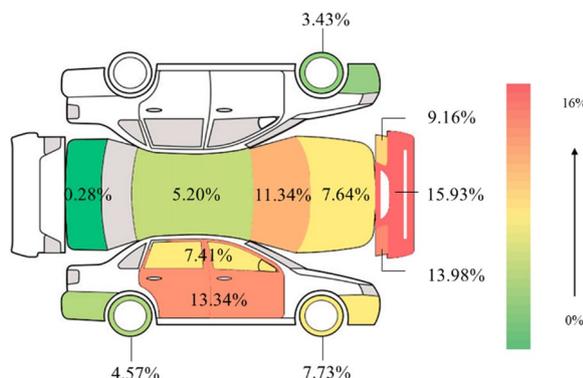
Fig. 13 Percentage of total gaze duration on each component

Fig. 14 Distribution of gaze duration across components as a function of the vehicle's distance

across each component. As shown in Fig. 13, the vehicle components that the participates mainly focused on were in the front (bumper, 15.93%; headlight, 23.14%) and the near side (near side door, 13.34%; side window, 7.41%) of a vehicle. In contrast, the gaze duration on the far side and back of the vehicle is shorter (far side front-wheel, 3.43%; luggage door, 0.28%).

Total gaze duration was related to the components and was affected by the position of vehicles. Figure 14 shows that the proportion of gaze duration on each component varied at different distances.

To explore how the fixation duration varied with components and distance, we recoded the 7956 fixations of 5750 vehicles into 6120 cases (10*3*17*12) according to four variables: longitudinal distance (10 groups), transverse distance (3 lanes), participant (17), and component (12). If two fixations were the same on all four variables, then we calculate the sum of their gaze duration. If a participant does not look at a component at a certain distance range, then we record the gaze duration as zero. Then we built a linear mixed model with component and longitudinal distance as fixed effects to verify the changes statistically. The core syntax of Model 3 is *Total gaze duration ~ Longitudinal distance * Component + (1| Participant).* The sign "*" means we consider both the main effect of each factor and the
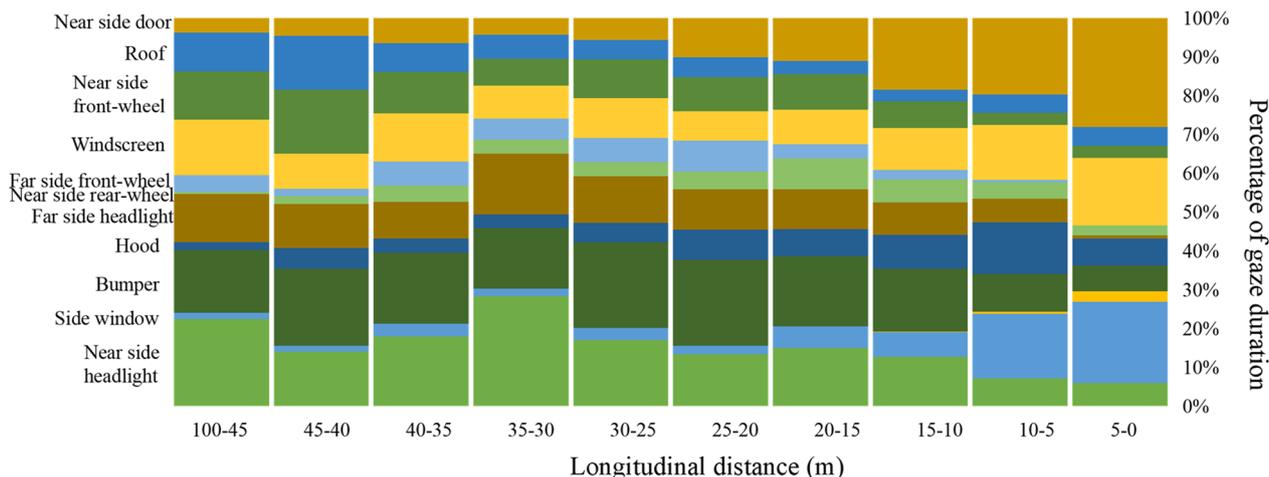
duration to a big vehicle than a normal-sized vehicle ($t = 2.44$, $p = 0.015$). Similarly, they spent more time gazing at a taxi than a private car ($t = 9.07$, $p < 0.001$).

4.3 Gaze pattern towards vehicle components

After displaying the effect of various features on gaze duration at the whole-vehicle level, we divided a vehicle into twelve components and explored the gaze duration

Table 5 The effect of vehicle component and longitudinal distance on gaze duration

	NumDF	DenDF	F value	p
Longitudinal distance	9	5984	42.47	< 0.001[***]
Component	11	5984	27.41	< 0.001[***]
Longitudinal distance* component	99	5984	3.83	< 0.001[***]

***$p < 0.001$

interaction between two factors. The estimation for the main effects and interaction effect are shown in Table 5.

The results showed that both longitudinal distance and components have significant effects on the total gaze duration. In addition, the interaction between distance and components was significant, which means pedestrian's gaze duration on the components varied with distances.

Next, we made pairwise comparisons of the gaze duration on components at each distance range to identify the main gaze area at each distance range. When the vehicle was far away (> 45 m), participants' gaze duration on the near side headlight was significantly longer than other components. With vehicles approaching (15–30 m), the gaze duration on the bumper increased and became the main focus area just like the near side headlight. When the vehicle was close to the participants (5–15 m), participants' main gaze area began to include the front and near sides of the vehicle, such as the far side headlight, hood, windscreen, near side door, and the near side window. Finally, when the distance between the vehicle and participants was within five meters, the only area where gaze duration was significantly longer was the near side door of vehicles.

5 Discussion

The objective of the study was to identify pedestrians' gaze patterns before crossing in a naturalistic setting. Our core finding was that pedestrians' fixation characteristics varied with the position of the gaze relative to the overall scene, the attributes of the gazed vehicles, and the components of a vehicle. In this section, we discussed how these moderators affected pedestrian gaze patterns, hoping to link the gaze pattern with the decision-making process of pedestrians and conclude theoretical and practical implications.

5.1 Gaze patterns moderated by vehicle features

Since vehicles are the main gazed target, we mainly discussed how vehicle features such as distance, size, and type affected pedestrian gaze patterns.

5.1.1 Vehicle distance

In the longitudinal direction, our results showed that vehicles' distance significantly affected pedestrians' gaze duration. Specifically, pedestrians gaze at vehicles at far distances (25–90 m) with low frequency but long duration. However, when the vehicles are between 25 and 5 m, pedestrian gazes towards them become more frequent yet shorter. One possible explanation is that pedestrians' gaze pattern is related to their information processing need to make a crossing decision and visible time of vehicles.

To make a road-crossing decision, pedestrians need to evaluate whether the gap between vehicles is large enough to cross without conflict with vehicles, often referred to as the gap acceptance process [24]. In this process, a gap too small or too large is intuitively rejected or accepted, thus does not require detailed evaluation. In contrast, vehicles with an arrival time of 2–8 s have a significant impact on pedestrians' gap acceptance [25]. Translating this arrival range of time, based on the speed range of vehicles in our experiment scene (30–50 km/h), the corresponding distance range is roughly 17–111 m. The distance range generally overlapped with the "low frequency but long duration" range that we found in our heat maps (25–90 m, Figs. 9, 10). Therefore, we speculate that pedestrians gazed towards far vehicles mainly to obtain information to support their road-crossing decision.

A previous study has found that most pedestrians rejected the gap smaller than 2 s [25]. Correspondingly, we found our participants made frequent but shorter gazes towards vehicles in this range (5–20 m). We speculate that pedestrians' frequent gaze is to monitor and check vehicles' state of motion. If vehicles continue to approach at a constant speed or even accelerate, pedestrians will keep their initial rejection decisions. However, if the vehicles show signs of slowing down or even giving way, the pedestrians may change the decision to use the crossing chance. At this stage, pedestrians have to interact more frequently with drivers, but checking the motion state is relatively simple. Also, the closer the distance between the vehicle and pedestrians, the shorter the time available for the pedestrian to observe, as the closer vehicles will drive past the pedestrian quickly and will no longer have an impact on the decision. Hence, pedestrians show a longer total but shorter mean gaze duration. However, these are all our speculation based on experience in this field and still require further empirical evidence to confirm.

In the transverse direction, pedestrians' gaze duration is differently distributed across different lanes. There is a clear trend of shorter gaze duration at a farther transversal distance. The trend implies that pedestrians allocate attention to vehicles based on the transverse distances from their standing position. This finding provided direct evidence on the "*rolling gap*" strategy while making gap acceptance decisions, which differs from classic assumptions of gap acceptance theory [24].

In the classic gap acceptance framework, pedestrians are assumed to look at all lanes and consider the gaps produced by the head-most vehicles. This assumption is inherited from previous studies on drivers' gap acceptance behavior when they arrive at intersections [26]. However, compared with drivers, pedestrian behaviors

are more flexible and adjustable. For example, pedestrians may cross a multi-lane street in stages because their speed is much lower than vehicles and their size is smaller. Thus, they can stop in the middle of a road to gather more information for further actions, while it is hard for vehicles to stop and observe at intersections. As a result, some previous researchers suggest that pedestrians may use different crossing strategies, such as rolling gaps [24, 27]. This strategy allows pedestrians to cross when an acceptable gap occurs in the adjacent lane, even if the gaps for the head-most vehicles (in farther lanes) are too small. Currently, the primary evidence of the rolling gap strategy is observational findings of pedestrian crossing behavior that conflicted with the classic assumption of gap acceptance. Therefore, our findings of on the less gaze duration at farther lanes can be the direct evidence of the rolling gap strategy.

5.1.2 Vehicle size and type
We found that compared with normal-sized vehicles, large vehicles were gazed longer. The longer gaze may reflect an attention bias towards more threatening objects. Consistent with this trend, previous studied reported that pedestrians judged the large vehicles as arriving earlier [28] and accepted larger critical gap [25, 29]. Therefore, we speculate that these cognitive differences between large and small vehicles may be related to the initial gaze phase.

In terms of types, gaze duration on the bus was the shortest, followed by regular cars, and the most prolonged duration is on taxis. Compared with the other types, taxis are the vehicles that interact most with pedestrians. Taxi drivers are more likely to slow down or even stop when to identify potential passengers waiting on the side of the road. Facing the communication signal sent by the taxi drivers, pedestrians are also likely to fixate more on taxies as responses. Previous studies also found that pedestrians' crossing decision varies between passenger cars and others (such as taxi and bus) [29]. It may be another reason why vehicle type significantly affects gaze duration.

5.2 Dynamic gaze pattern towards vehicles' components
In general, as vehicles approaching, pedestrian's gaze areas expand from the near side headlight to the whole front and near side, finally shift to the near side (near side door and near side window). We assume this gaze pattern may be explained by the limitations in view angle range. When the vehicle is far in the natural multi-lane road, the near side headlight and bumper are the most easily observed areas. When the distance is less than five meters, or even directly in front of the pedestrians, the near side door is the largest and most apparent area in their field of vision. Therefore, the visibility of a component in the field of view may be one of the reasons affecting the gazed components.

5.3 Implications
Our study contributes to both theory and application. As the first study recorded pedestrian gaze behavior during crossing decision-making in the natural environment, it intuitively demonstrated the cues and strategies (i.e., rolling gap) employed by pedestrians during decision-making. What is more, the distribution of pedestrian gaze across different positions and features defined the potential zone of interaction between pedestrians and vehicles, which can help build realistic assumptions on whether a vehicle needs to be considered while modeling pedestrian behavior [30]. As an example, we found that participants paid less attention to the farther lanes. Then a moderator can be added to attenuate the role of vehicles in farther lanes while modeling pedestrian risk perception from vehicles.

In practice, the above interaction zone has design indications on the communication between autonomous vehicles and pedestrians. A previous study reported that the location and timing of information display on external human–machine interfaces (eHMIs) significantly affect pedestrians' safety feelings and crossing willingness [14, 19]. Therefore, based on the user-centered design principles, we suggest displaying the dynamic eHMIs with pedestrians' gaze patterns. Specifically, we suggest displaying interaction information in the near side headlight when the vehicles are 30 m away. The bumper can also be one of the display options between 30 and 15 m. From 15 to 5 m, the whole front and near side body of vehicles can be used to place the eHMIs, but closer than 5 m, the near side body becomes the only suitable choice for display.

5.4 Limitations
There are some limitations to consider while interpreting the findings. First, we coded the fixations manually which may introduce ambiguity in the judgment of longitudinal distance and vehicle components. For the participants, although our sample size (N=17) is common in studies used a similar approach [10, 13], more accurate and quantitative portrait of the gaze pattern may require more participants with various ages. Also, despite our best efforts to control factors such as weather, time period, and so on, we recognize that we cannot guarantee that each participant will face the same traffic environment.

Second, our raw data were fixation trajectories of pedestrian gaze towards specific targets and a video recording of their surrounding scenes. When pedestrian

fixations jumped from one target to another, it is challenging to build a continuous image of the context of pedestrian gaze behavior. Therefore, the current findings can only be viewed as a coarse-grained profile of the holistic gaze pattern without contextual information. A related consequence is that we cannot extract detailed data continuously for a target. For example, although vehicles' time to arrival is important in determining gaze behavior, we cannot extract this variable because the vehicle may be invisible once pedestrian changed their fixation. As a result, we can only estimate the time to arrival of vehicles based on the general speed of vehicles in the discussion. For future research, we recommend measuring instant speed in a less complex site to distinguish the role of distance and vehicles' time-to-arrival (TTA) in determining the gaze pattern.

Third, although we asked participants to simulate the real process of waiting to cross road, we cannot distinguish whether they could do it. A participant waiting for a simulated cross and a pedestrian waiting to cross may yield different gaze patterns. To exclude conditions where participants rush towards the destination, all participants were told there was no time limit, and they could wait until they felt it was safe as in daily life when they were not in a hurry. However, in real-life scenarios, pedestrian decisions may be affected by their state and other tasks [31]. To extend the current findings in a general situation, subsequent studies can include more factors related to pedestrian state (e.g., distracted, accompanied).

Finally, we observed the gaze behavior of pedestrians using an eye-tracker and had obtained some descriptive conclusions, but the mechanism of these patterns is still unknown. For example, what is the relationship between fixations and attention? What cognitive processes do the average and total gaze duration reflect? These issues call for experimental studies with more detailed measurement of pedestrian psychological state.

6 Conclusions

In this study, we used the eye-tracker to record pedestrians' gaze behavior before crossing the road and analyzed pedestrian gaze duration towards the overall scene, vehicles, and components. We found that pedestrians' gaze towards vehicles at far distances (25–90 m) is less frequent but longer in the longitudinal direction. When the vehicles are between 25 and 5 m, pedestrians' fixations are more frequent yet shorter. In addition, we found a significant difference in pedestrians' gaze duration between lanes in the transverse direction, which supports the "rolling gap" strategy. For a specific vehicle, large vehicles and taxis were gazed longer compared with normal vehicles. Pedestrians also show a dynamic gaze pattern towards vehicles' components, which gradually

changes from the near side front (near side headlight) to the whole near side (headlight, bumper, hood, and windscreen), and finally, shift to the near side (near side door and near side window) as the vehicle approaches. Our study contributes to understanding pedestrians' cognitive processes and gaze behavior characteristics while making road-crossing decisions. In practice, these findings can also provide reference comments and design inspiration on the design of eHMIs in automated vehicles.

Author contributions

QZ: Conceptualization, Data collection, Writing, Formal analysis. XZ: Conceptualization, Writing, Review and editing. TZ: Data collection and analysis. YH: Data collection and analysis. GM: Conceptualization, Review and editing.

Availability of data and materials

Please contact the author for the original material and data.

Declarations

Competing interests

The authors declare that they have no competing of interest.

References

1. WHO. (2018). Global status report on road safety 2018 [R/OL]. [2020-10-04]. https://apps.who.int/iris/handle/10665/276462.
2. Lassarre, S., Papadimitriou, E., Yannis, G., & Golias, J. (2007). Measuring accident risk exposure for pedestrians in different micro-environments. *Accident Analysis & Prevention, 39*(6), 1226–1238. https://doi.org/10.1016/j.aap.2007.03.009
3. Older, S. J., & Grayson, G. B. (1974). *Perception and decision in the pedestrian task.* Transport and Road Research Laboratory.
4. Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37,* 32–64. https://doi.org/10.1518/001872095779049543
5. Jacob, R. J. K. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems (TOIS), 9*(2), 152–169. https://doi.org/10.1145/123078.128728
6. Clercq, K. D., Dietrich, A., Velasco, J. P. N., Winter, J. D., & Happee, R. (2019). External human-machine interfaces on automated vehicles: Effects on

Zhao *et al. European Transport Research Review*   (2023) 15:31

Page 16 of 16

pedestrian crossing decisions. *Human Factors, 61*(8), 1353–1370. https://doi.org/10.1177/0018720819836343

7. Weber, F., Chadowitz, R., Schmidt, K., Messerschmidt, J., & Fuest, T. (2019). Crossing the street across the globe: A Study on the effects of eHMI on pedestrians in the US, Germany and China. In H. Krömker (Ed.), *HCI in mobility, transport, and automotive systems* (pp. 515–530). Springer.

8. Hollands, M., Marple-Horvat, D. E., Henkes, S., & Rowan, A. K. (1995). Human eye movement during visually guided stepping. *Journal of Motor Behavior, 27*, 155–163. https://doi.org/10.1080/00222895.1995.9941707

9. Marigold, D. S., & Patla, A. E. (2007). Gaze fixation patterns for negotiating complex ground terrain. *Neuroscience, 144*(1), 302–313. https://doi.org/10.1016/j.neuroscience.2006.09.006

10. Hollands, M., Patla, A. E., & Vickers, J. N. (2002). "Look where you're going!": Gaze behaviour associated with maintaining and changing the direction of locomotion. *Experimental Brain Research, 143*(2), 221–230. https://doi.org/10.1007/s00221-001-0983-7

11. Thomson, J. A., Tolmie, A., Foot, H. C. & McLaren, B. (1996). Child development and the aims of road safety education: A review and analysis. Department of Transport Road Safety Research Report No. 1. London: HMSO.

12. Trefzger, M., Blascheck, T., Raschke, M., Hausmann, S., Schlegel, T. (2018). A visual comparison of gaze behavior from pedestrians and cyclists. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, Warsaw, Poland, 14–17.

13. Geruschat, D., Hassan, S., & Turano, K. (2003). Gaze behavior while crossing complex intersections. *American Academy of Optometry, 80*(7), 515–528. https://doi.org/10.1097/00006324-200307000-00013

14. Dey, D., Walker, F., Martens, M., & Terken, J. (2019). Gaze patterns in pedestrian interaction with vehicles: towards effective design of external human-machine interfaces for automated vehicles. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (New York, NY: ACM), 369–378.

15. Leibowitz, H. W. (1985). Grade crossing accidents and human factors engineering: How a discipline combining technology and behavioral science can help reduce traffic fatalities. *American Scientist, 73*(6), 558–562.

16. Clark, H. E., Perrone, J. A., Isler, R. B., & Charlton, S. G. (2016). The role of eye movements in the size-speed illusion of approaching trains. *Accident Analysis and Prevention, 86*, 146–154. https://doi.org/10.1016/j.aap.2015.10.028

17. Zhang, N. (2012). Jiyu shenjing renyinxue de ren-che jiaohu xitong ruogan wenti de yanjiu [Research on several issues in human-vehicle interaction system based on neuroergonomics]. (Doctoral Dissertation). Northeastern University. https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFDLAST2015&filename=1015562191.nh.

18. Newstead, S., & D'Elia, A. (2010). Does vehicle colour influence crash risk? *Safety Science, 48*(10), 1327–1338. https://doi.org/10.1016/j.ssci.2010.05.001

19. Eisma, Y. B., Bergen, S. V., Brake, S. M. T., Hensen, M. T. T., Tempelaar, W. J., & Winter, J. C. F. D. (2020). External human–machine interfaces: The effect of display location on crossing intentions and eye movements. *Information, 11*(1), 13. https://doi.org/10.3390/info11010013

20. Ackermann, C., Beggiato, M., Schubert, S., & Krems, J. F. (2019). An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles? *Applied Ergonomics, 75*, 272–282. https://doi.org/10.1016/j.apergo.2018.11.002

21. Winter, J. D., Bazilinskyy, P., Wesdorp, D., Vlam, V. D., Hopmans, B., Visscher, J., & Dodou, D. (2020). How do pedestrians distribute their visual attention when walking through a parking garage? An eye-tracking study. *Ergonomics, 64*(6), 793–805. https://doi.org/10.1080/00140139.2020.1862310

22. Himanen, V., & Kulmala, R. (1988). An application of logit models in analysing the behaviour of pedestrians and car drivers on pedestrian crossings. *Accident Analysis & Prevention, 20*(3), 187–197. https://doi.org/10.1016/0001-4575(88)90003-6

23. Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

24. Brewer, M., Fitzpatrick, K., Whitacre, J. A., & Lord, D. (2006). Exploration of pedestrian gap-acceptance behavior at selected locations. *Transportation Research Record, 1982*(1), 132–140. https://doi.org/10.1177/0361198106198200117

25. Das, S., Manski, C. F., & Manuszak, M. D. (2005). Walk or wait? An empirical analysis of street crossing decisions. *Journal of Applied Econometrics, 20*(4), 529–548. https://doi.org/10.1002/jae.791

26. Brilon, W., Koenig, R., & Troutbeck, R. J. (1999). Useful estimation procedures for critical gaps. *Transportation Research Part A: Policy & Practice, 33*(3/4), 161–186. https://doi.org/10.1016/S0965-8564(98)00048-2

27. Zhuang, X., & Wu, C. (2011). Pedestrians' crossing behaviors and safety at unmarked roadway in China. *Accident Analysis & Prevention, 43*(6), 1927–1936.

28. Petzoldt, T., Hoang, Q., & Bogda, K. (2017). Time to arrival estimates, (pedestrian) gap acceptance and the size arrival effect. In *Proceedings of the 9th International Driving Symposium on Human Factors in Driving Assessment, Training, and Vehicle Design*. Manchester, Village, Vermont.

29. Yannis, G., Papadimitriou, E., & Theofilatos, A. (2013). Pedestrian gap acceptance for mid-block street crossing. *Transportation Planning and Technology, 36*(5), 450–462. https://doi.org/10.1080/03081060.2013.818274

30. Zhuang, X., & Wu, C. (2013). Modeling pedestrian crossing paths at unmarked roadways. *IEEE Transactions on Intelligent Transportation Systems, 14*, 1438–1448. https://doi.org/10.1109/TITS.2013.2267734

31. Zhuang, X., Wu, C., & Ma, S. (2018). Cross or wait? Pedestrian decision-making during clearance phase at signalized intersections. *Accident Analysis Prevention, 111*, 115–124. https://doi.org/10.1016/j.aap.2017.08.019

Publisher's Note