Where Do Pedestrians Look When Crossing? A State of the Art of the Eye-Tracking Studies

LUCIE LÉVÊQUE, MAUD RANCHET, JONATHAN DENIEL, JEAN-CHARLES BORNARD, AND THIERRY BELLET

1Laboratory of Ergonomics and Cognitive Sciences Applied to Transport, Gustave Eiffel University, 69675 Lyon, France
2CIVITEC, Engineering Systems International (ESI) Group, 69458 Lyon, France

Corresponding author: Lucie Lévêque (lucie.leveque@univ-eiffel.fr).

This work was supported in part by the European Union’s Horizon 2020 Research and Innovation Program (i.e., SUaaVE project: SUpporting acceptance of automatedVEhicles), under Grant 814999.

ABSTRACT It has been widely shown in the literature that analysing eye movements and positions can provide useful information for a better understanding of human perception and cognition. The eye-tracking technology, as a process of measuring where people look, has established itself as a widespread means of studying visual information processing in several domains, including in the study of human walking. Street-crossing can be defined as a particular form of walking. Indeed, several elements have to be considered in the decision-making process, such as the distance headway, traffic density, vehicle speed, etc. It is also a very risky aspect of walking as pedestrians are considered one of the most vulnerable road users. In this article, we present an up-to-date comprehensive review of existing eye-tracking experiments in the literature, from the pedestrian’s point of view, with a view to study the effects of both internal (e.g., age) and external (e.g., road environment) factors on pedestrians’ road crossing gaze behaviour. Furthermore, the current gaps in the literature are then discussed in order to open up some future perspectives in the field, such as the forthcoming introduction of automated vehicles on the roads.

INDEX TERMS Eye-tracking, visual attention, street-crossing, pedestrians, automated vehicles.

I. INTRODUCTION

The eye-tracking technology, that is to say the process of measuring where humans look in a visual field, has become a very useful tool in various domains including, but not limited to, psychology, usability, and human computer interaction (HMI) [1]–[4]. The eye-tracking technology generally allows the study of human visual attention allocation, which aims to better understand how the cognitive system processes the overloading flow of information it receives in a continuous manner. This is driven by both bottom-up and top-down attentional processes [5]–[7]. More precisely, bottom-up aspects are based on a visual scene’s characteristics, whereas top-down attention is determined by other factors like knowledge, experience, intentions, expectations, and schemas [8], [9]. Therefore, tracking eye movements can provide information to better understand perception, allowing a glance into cognition.

Very first attempts to observe and even record eye movements and gaze behaviour can be traced back to Javal’s research, at the end of the nineteenth century. Back then, everything had yet to be discovered on the topic; from the basic description and taxonomy of eye movements to observation and measurement tools and techniques. The first half of the twentieth century was marked by the technical development of prototypic tools for eye movements detection and tracking. These tools were usually very invasive and provided inaccurate quality measurements. However, they still allowed the classification of most of the eye movements that are now commonly used in eye-tracking studies. Later, in the 1950s, Fitts et al. [10] used the eye-tracking process to study pilots’ gaze behaviour and improve cockpits’ instrument clusters arrangement. By that time, eye-tracking data collection was extremely tedious and most of technological issues still had to be overcome, but the concept was already acknowledged as promising. The second half of the twentieth century saw eye-tracking systems raise from prototypic tools to market products. Several technical issues were addressed, such as measure accuracy or device portability.

The associate editor coordinating the review of this manuscript and approving it for publication was Tomasz Trzcinski.

Received August 25, 2020, accepted August 30, 2020, date of publication September 2, 2020, date of current version September 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3021208
(for instance, Young et al. [11] provided photographic examples of several eye-tracking systems). Data processing was also an important issue, and this question became tightly related to computer technology developments, as provided data still often required to be manually processed [12]. In addition, metrics had to be developed in parallel of the technology improvements to infer what aspects of gaze behaviour are relevant to assess subjects’ ongoing cognitive processes [13].

After a century of advances in the aforementioned technology, mobile eye-trackers opened up possibilities to explore eye movements during dynamic activities, such as walking. Walking, especially in unfamiliar or complex environments, requires pedestrians to explore the visual scene, process relevant information, and make decisions. During locomotion, results clearly showed that combined yaw, pitch, and roll movements of the body, head, and eye maintain the gaze in the direction of forward motion during straight walking [14]–[16]. Acquiring information about a visual scene is then achieved thanks to a combination of eye fixations and saccades. Although various research investigated eye movements and positions before and while walking, as far as we know only few studies explored the precise nature of visual information used in specific walking situations such as crossing a street.

Street-crossing involves taking into account multiple considerations. Indeed, before a pedestrian initiates the crossing, they have to make a decision based on their perception of safety. According to several authors, including Wang et al. [17] and Sucha et al. [18], a pedestrian mainly makes their crossing decision when the gap size, i.e., distance between two cars or distance between the pedestrian and next car, is large enough. Such a gap obviously depends on the vehicle speed; the higher the speed, the larger the gap has to be. Gap estimation is a complex task for pedestrians, as a bigger object seems to approach earlier than a smaller one, for instance. Interestingly, pedestrians seem to show a better speed estimation when a car is approaching from the opposite direction of their walking direction [19]. Other factors have to be considered in the decision-making, like the traffic density (i.e., the number of vehicles occupying a unit length of roadway). Eventually, pedestrians present diverse inner characteristics (e.g., age and gender), potentially leading to different crossing behaviours.

In this article, a comprehensive literature review of eye-tracking studies focusing on road crossing situations is presented. Existing surveys found in the literature focus on the driver’s point of view [20], [21], whereas our article is centred on the pedestrian’s point of view. The effects of internal and external factors on pedestrians’ crossing behaviour will be discussed in section II and III, respectively. In section IV, the eye-tracking studies presented, as well as their limitations, will be discussed through different yet complementary approaches. Such research work aims to reach several objectives, e.g., to better understand and consequently improve road traffic safety.

II. EFFECTS OF INTERNAL FACTORS ON THE GAZE BEHAVIOUR OF PEDESTRIANS BEFORE AND WHILE CROSSING

Table 1 summarises existing eye-tracking studies which investigated whether, and to what extent, internal factors affect pedestrians’ gaze behaviours before and while crossing. To the best of our knowledge, age is the only intrinsic factor studied in the literature to date. The following two sub-sections detail the works comparing children and adults, and younger and older adults, respectively.

A. COMPARISON BETWEEN CHILDREN AND ADULTS

Egan et al. seem to be the precursors in the use of eye-tracking technology to compare the visual search strategies of children and adults in road crossing [22]. Indeed, in a study published in 2008, the authors explored the eye movements of ten children, aged 8 to 9 years old, and ten university students, aged 20 to 28 (i.e., reckoned adults), while crossing diverse roads using a portable ASL (Applied Science Laboratories) Mobile Eye device. Figure 1 illustrates diverse types of eye-tracking devices used in the literature; the ASL Mobile Eye belongs to the glasses’ category (i.e., Figure 1 (c)). The participants were asked to cross a total of four to six signalised crossings (i.e., consisting of pedestrian traffic lights), as well as several unsignalised crossings, around their own school or university (depending on the group studied). Two main features were analysed, namely the gaze position (i.e., on different regions of interest) and the gaze direction (i.e., left, centre, or right) of the participants. Findings showed that adults significantly spent more time looking at the road and at vehicles than children did, and this both under signalised and unsignalised crossings conditions.

Children happened to fixate on irrelevant objects, such as pedestrians or buildings. Moreover, results showed that adults looked more to the right, i.e., towards the direction...
TABLE 1. Overview of the eye-tracking studies investigating the impact of internal factors on street-crossing gaze behaviour.

| Source article | Internal factor studied | Participants | Method | Eye-tracking device | Gaze metrics | Main findings |
|----------------|-------------------------|--------------|--------|---------------------|--------------|---------------|
| Biassoni et al. [22] | Age | 22 adults (20-65 years old), 22 pre-school children (5-6 years old) | 4 pictures of pedestrian crossings (3 with zebras, 1 with traffic light), 5 s each | Screen-based Tobii X120 (60 Hz) | Fixation length and fixation count for each of the 6 ROIs defined | Adults used a dynamic exploration behaviour (they looked more frequently and longer at the ROIs), children looked at irrelevant but salient elements |
| Egan et al. [23]-[24] | Age | 10 students (20-28 years old), 10 children (8-9 years old) | 4-6 outdoor signalised crossings, various unsignalised crossings | ASL Mobile Eye | Gaze position, gaze direction | Adults looked more at the road and vehicles, and more to the direction of oncoming traffic, than children who fixated on irrelevant information |
| Tapiro et al. [25] | Age | 21 adults (20-27 years old), 33 children (7-13 years old, in 3 age groups) | 18 scenarios in mixed reality simulator (3D Perception™ Dome), 10-45 s each | Head-mounted ASL HS-H6 (60 Hz) | Total gaze distribution, fixation duration, dwells, frequency of fixations | Adults shifted their attention to the far left, whereas children spent more time looking at the central area |
| Tapiro et al. [26] | Age | 11 students (25-30 years old), 10 older adults (over 65 years old) | 6 simulated scenarios (3D Perception™ Dome), 60-90 s each | Ergoneers Dikablis glasses (60 Hz) | Total gaze distribution, transition matrix (probabilities of shifting gaze from a ROI to another) | Younger adults spent more time looking at the left and right ROIs than the older ones, who allocated more attention to the centre |
| Tapiro et al. [27] | Age | 51 adults (22-29 years old), 32 children (9-13 years old, in 2 age groups) | 12 simulated scenarios (3D Perception™ Dome), 60 s each | ASL | Visual attention dispersion | Children showed higher dispersion than adults when exposed to higher visual load conditions |
| Zito et al. [28] | Age | 20 adults (23-28 years old), 20 older adults (65-79 years old) | 2 simulated scenarios (Först GmbH) | Head-mounted SMI iView X (50-60 Hz) | Number of fixations and head movements for each of the 3 ROIs defined | Similar exploration patterns for both groups with a larger focus on the left, but younger group explored the right part more often |

Abbreviations used in the table: ASL: Applied Science Laboratories, ROI: Region of Interest.

of oncoming traffic, than children did. The (lucky) children who participated in this first study were then rewarded with an ice cream, obviously with a view to support scientific advances [23]. The ice cream was indeed used as a distractor in a second phase of the experiment, where the children were asked to cross the roads once again. Note that adults did not take part in this second experiment. During the distractor condition, i.e., with ice cream, the children focused even less on their environment and more in the centre than on the left and right parts compared with the control condition, that is to say, without any ice cream.

In a study published in 2013, Tapiro et al. also compared the visual search strategies of children and adults in road-crossing situations [24]. Fifty-four participants were
recruited, including twenty-one adults (aged 20 to 27), fourteen children aged 7 to 8 years old, eighteen children aged 9 to 10, and nine children aged 11 to 13. All the participants were asked to watch eighteen urban road-crossing scenarios through a simulator of virtual reality (VR), the 3D Perception™ Dome (as shown in Figure 3 (c)); and to press a button when the situation was perceived as safe to cross. Their eye movements were recorded using a head-mounted ASL HS-H6 eye-tracker. Through the analysis of four measures, i.e., gaze distribution, fixation duration, dwells, and frequency of fixations, the authors showed significant differences in terms of visual behaviour between children and adults. Indeed, children spent more time looking at the closest region of interest (ROI), whereas adults spent more time looking at the far-left range (i.e., at cars traveling on the closest lane). Average fixation duration seemed to increase with age, as youngest children showed the shortest fixations and adults showed the longest ones. Similarly, adults displayed longer dwellings in peripheral areas. Finally, youngest children had the highest fixation frequency. It is interesting to note that, as a whole, the oldest group of children exhibited visual strategies resembling that of adults.

Biassoni et al. also compared the ocular behaviour of children and adults in crossing situations using an eye-tracking device [22]. In their study, participants were instructed to observe four pictures of pedestrian crossings (Figure 3 (b)) presented for five seconds each on a computer screen. More precisely, they were asked to observe them as if they were on the sidewalk and had to cross the road. Three pictures represented two different pedestrian crossings without traffic lights and very low traffic conditions, whereas the fourth one showed a pedestrian crossing regulated by traffic lights. The pictures were taken both from an adult’s and a child’s point of view. To analyse the ocular behaviour of the participants, six regions of interest (ROIs) were created; the four areas around the zebra crossing were similar between pictures. Mean values of fixation length, as well as mean values of fixation count, were compared between the two groups. Findings consistently showed that adults looked more frequently, and for a longer time, at all the different areas of the field of view. Furthermore, their gaze shifted from one area to another. Conversely, children tended to focus on a few areas on smaller parts of the field of view. Another result is that adults looked at different areas than children. Indeed, adults particularly scanned distant areas where a vehicle could suddenly approach the zebra crossing, which was not the case of children. The latter looked at elements considered irrelevant to cross the road safely (e.g., stopped tramway and parked vehicles).

Another study focusing on differences between children and adults’ visual attention was conducted in 2020 by Tapiro et al. [27]. In this recent study, the authors aimed to examine whether, and to what extent, urban street facade affects adults and children’s gaze behaviour when crossing. Eighty-three participants were recruited, composed of fifty-one adults (aged 22 to 29), twenty-one children aged 9 to 10, and eleven children aged 11 to 13. The same simulator than in their previous study [25], i.e., 3D Perception™ Dome, was used for this study, where participants were shown twelve scenarios without crosswalk. Again, participants had to indicate their crossing wish by pressing a button. Their visual attention dispersion, i.e., standard deviations of the fixations from the scene centre, was measured thanks to an ASL eye-tracking device (note no further detail is given by the authors regarding the characteristics of the eye-tracker). Once more, results analysis showed discrepancies between the age groups. For instance, children’s gaze behaviour was
more dispersed than that of adults. More results for this study are given in section III as far as the impact of the road environment on visual search is concerned.

B. COMPARISON BETWEEN YOUNGER AND OLDER ADULTS

Tapiro et al. also conducted an eye-tracking study on the same theme, i.e., on the impact of age on road-crossing gaze behaviour, this time with older participants [26]. Indeed, in 2016, the authors investigated how older pedestrians spread their visual attention before crossing. Twenty-one participants were involved in the experiment, including eleven university students (aged 25 to 30 years old), and ten adults aged over 65 years old. Once again, pedestrian simulator using virtual reality was used in order to display six scenarios, lasting between sixty and ninety seconds each. In particular, scenarios represented a two-way traffic flow surrounded by urban features (e.g., street signs and bus stops). The participants had to press a button as soon as they made their decision to cross the street. It can be noted that there was, at least, one crossing opportunity (i.e., sufficient gap between two cars) in each scenario. Ergoneers Dikablis glasses were used to record participants’ eye movements; more specifically, the gaze distribution and transition matrix were analysed. The results showed that older and young adults differently divided their attention on the ROIs. Indeed, the older group spent more time focusing on the central area of the scene than the younger group. Also, the analyses showed that the older adults were less inclined to shift their gaze from the centre towards the side areas than the younger pedestrians. The authors suggested that the former may be too focused on their travel path.

Similarly, the study of Zito et al. aimed to better understand the decision-making process during street-crossing [27]. The novelty of this study was to measure behavioural data such as eye and head movements during street-crossing, as well as to investigate the differences between young adults and healthy older adults regarding the number of road crossings, virtual crashes, and missed crossing opportunities. After performing visual, cognitive, as well as a ten-metre gait speed tests, participants were instructed to achieve a street-crossing task using a modified version of a driving simulator. They were standing in front of a projected crosswalk of a two-way road while cars were driving in the nearest lane, i.e., from left to right. Participants were presented two different scenarios where six cars were driving at the same, constant speed, with a gap varying in time between cars. The participants had to indicate their decision to cross the street by taking a step forward. In total, thirty trials, with different time gaps and car speeds, were presented in a randomised order. During the street-crossing task, the number of visual fixations on three regions of interests (i.e., left and right parts of the screen, and floor below the screen) was measured. Head-tracking outcome, that is to say the number of head movements rightward, leftward, and downward, was also measured. Findings showed that the percentage of fixations towards the left part of the screen was very high for both groups. However, older people looked less at the other side of the street to make their crossing decision, and they had a higher percentage of fixations below the screen compared to the younger ones. Furthermore, they overestimated their walking speed. The authors concluded that older adults had more difficulties than young adults in making the decision to cross, especially under time pressure.

III. EFFECTS OF EXTERNAL FACTORS ON THE GAZE BEHAVIOUR OF PEDESTRIANS BEFORE AND WHILE CROSSING

Table 2 is a summary of the eye-tracking studies which investigated the effects of external factors on pedestrians’ behaviours before and while crossing.

A. EFFECTS OF THE ROAD ENVIRONMENT

One of the pioneering works investigating visual exploration strategies of pedestrians while crossing was conducted by Geruschat et al. in 2003 [31]. Their goal was to assess how normally sighted people use their vision to cross a street safely. Twelve participants were recruited including three young adults and nine older adults. Participants were equipped with a head-mounted Iscan ETL-500 eye-tracking device, and instructed to cross back and forth two different types of intersections as they would in real life. The first intersection was a plus shaped cross with traffic lights; the second one was a roundabout exit with no traffic lights. Figure 2 exemplifies different types of intersections for a better understanding of the situations presented. Results showed that the type of intersection had a significant influence on visual strategies adopted by the pedestrians. Firstly, when participants were waiting at the curb, the type of intersection had an impact on their exploration strategy. Indeed, when there were no traffic lights, participants tended to look for potential incoming traffic to detect a safe gap and decide to cross. On the contrary, with traffic lights, several participants relied on them to decide when to cross, while some others kept using the same gap detection strategy as detailed previously. Secondly, when the participants were crossing, the type of intersection influenced whether or not they would redirect their gaze towards the right. In fact, when there were two lanes to cross, participants looked for cars coming from the right; whereas, in the case of unidirectional traffic, participants redirected their gaze towards their ahead destination goal. The authors found common head and eye behaviours near the critical moments of crossing the street.

In the study conducted by Tapiro et al., already presented in section II, three principal external elements were also investigated, namely: the traffic movement (no vehicle, one direction, two directions), the field of view (unrestricted vs. partially obscured), and the presence or absence of zebra-crossing [25]. Traffic movement had a significant impact on participants, i.e., they were more attentive to the road in the presence of driving vehicles than in their absence. Similarly, changes in the field of view led to significant
TABLE 2. Overview of the eye-tracking studies investigating the impact of external factors on street-crossing gaze behaviour.

| Source article | External factor(s) studied | Participants | Method | Eye-tracking device | Gaze metrics | Main findings |
|----------------|---------------------------|--------------|--------|--------------------|--------------|---------------|
| Dey et al. [29] | Traffic                   | 26 adults (11 male, 15 female) | 8 outdoor scenarios (4 yielding, 4 non-yielding behaviours) | Tobii Pro Glasses 2 (50 Hz) | Number of fixations and frequency of fixations for each of the 7 ROIs defined | First fixations on road surface, then on car’s bumper, hood, and windshield |
| Eisma et al. [30] | External human-machine interface (eHMI) | 61 students (51 male, 10 female) | 6 eHMI conditions (roof, windscreen, grill, projection, wheels, no eHMI), 6 environments | EyeLink 1000 Plus (2000 Hz) | Gaze spread | Concentrated gaze pattern on windscreen eHMI, dispersed pattern on projection eHMI |
| Geruschat et al. [31] | Intersection type | 12 adults | 2 intersections (light-controlled plus-type, roundabout without light) | Head-mounted Iscan ET5-500 (60 Hz) | Number of fixations, head direction | Absence of traffic lights leads to gazing towards oncoming cars |
| Jiang et al. [32] | Phone use | 28 college students (17 male, 11 female) | 4 outdoor scenarios (no distraction, music, phoning, texting) | Tobii Pro Glasses 2 (50 Hz) | Pupil diameter, scanning frequency, fixation points, time, duration for each of the 5 ROIs defined | Texting showed highest impairment, then phoning and music |
| Tapiro et al. [25] | Traffic, field of view, zebra-crossing | 21 adults and 33 children | 18 scenarios in mixed reality simulator (3D Perception™ Dome), 10-45 s each | Head-mounted ASL HS-H6 (60 Hz) | Total gaze distribution, fixation duration, dwells, frequency of fixations | Users more attentive when driving vehicles are present |
| Tapiro et al. [27] | Environment visual load | 51 adults and 32 children | 12 simulated scenarios (3D Perception™ Dome), 60 s each | ASL | Visual attention dispersion | Children showed higher dispersion than adults when exposed to higher visual load conditions |

Abbreviations used in the table: ASL: Applied Science Laboratories, ROI: Region of Interest.

differences; participants spent more time looking at areas closer to them when their field was restricted by parked vehicles. The authors did not specify results obtained regarding the diverse zebra-crossing conditions. As mentioned in the previous section, the same research team, i.e., Tapiro et al. recently studied how characteristics of the environment can impact the crossing behaviour [27]. For their experiment, three different levels of clutter (i.e., low load, medium load, and high load) were created, depending on the number of visual objects added to the simulated environment. This resulted in a higher rate of missed crossing opportunities in the case of a high level of visual clutter. However, when encountering such a high visual load, only the group of children presented more dispersed gazes.

B. EFFECTS OF NEW TECHNOLOGIES

Jiang et al. addressed the issue of mobile phones’ distraction on pedestrians while crossing [32]. They conducted an outdoor experiment where the type of phone usage was controlled. Twenty-eight college students were involved in the
study, where they had to perform eight road crossings around a plus-shaped intersection with traffic signals (as illustrated in Figure 2 (a)). Five of the crossings did not involve the use of mobile phone (note four of them were not considered in the study). For the three remaining ones, participants were asked to use their phone in different ways, namely: music listening, phone conversation, and texting. Participants were instructed to wait for the green light before crossing. The authors collected several data concerning crossing behaviour (i.e., time to initiate crossing, looks to the left and right, and crossing speed), and visual attention behaviour (i.e., fixation points, time, and duration, pupil diameter, and scanning frequency). Results showed that the time to initiate crossing was significantly lower under the undistracted condition compared to all three distracted ones. Texting appeared to have the most important impact on the time to initiate crossing, and it was associated with significantly less looks to the left- and right-hand sides. Crossing speed was also impacted by the type of distraction, with the lowest speeds for phone conversation and texting. Scanning frequency was negatively impacted by all the distraction conditions. Similarly, pupil diameter increased when participants were distracted. The fixation number, time, and mean duration were all affected by the use of mobile phone for each of the five ROIs defined: the more the task involved the participants to look at their phone, the less they looked at traffic, signals, and crosswalks.

Concerned by potential accidents between pedestrians and future automated vehicles, Dey et al. studied where pedestrians look before deciding to cross a road in the presence of an incoming car [29]. Their aim was to provide information about where, on automated vehicles, external human-machine interfaces (eHMIs) should display information to pedestrians. Twenty-six young adults were involved in the experiment. They were placed at the curb of a straight road with no zebras and only a single car, as in Figure 3 (a). Participants were instructed to proceed as if they had to cross and, as the car was approaching, to indicate their willingness to cross in real time through a potentiometer. The car yielded the right of way for half the trials. The authors found that, as long as the car was too far to represent any threat, participants were willing to cross and were mainly looking at the road area in front of the car. As the car approached, even if its speed was constantly decreasing, the participants’ intent to cross progressively decreased to reach a minimum when the distance to the participant ranged from thirty to fifteen metres. In this specific phase, participants’ gaze progressively switched from the road area to the car’s bumper and hood. When the car was between fifteen to five metres of the participant (i.e., stopped position to yield the right of way), their willingness to cross exponentially raised to maximum and was associated with a significant shift of the gaze towards the windshield. The authors assumed that participants were looking for cues from the driver to confirm their intention to yield.

Eisma et al. were also interested on the eHMIs which will potentially be placed on future automated (i.e., driverless) vehicles [30]. They carried out an eye-tracking experiment in order to investigate where pedestrians would look at on automated vehicles, with a view to provide more specific information on where to position such external interfaces. Sixty-one university students were asked to watch thirty-six video clips, with three different environments (i.e., straight road, T-junction, and cross-shaped intersection, which can be found in Figure 2), six different eHMI conditions, and two distinct traffic behaviours. More specifically, the external human-machine interfaces could be placed either: on the roof, windscreen, grill, wheels of the car, or as a projection on the road. The sixth eHMI condition was a controlled condition, i.e., without any eHMI involved. Each interface consisted of a text, either ‘driving’ or ‘waiting’, as well as an icon, to describe the car’s behaviour. At the end of each of the thirty-six scenarios, the participants had to rate the following statement: “it was clear when I could cross”, from 0 (i.e., completely disagree) to 10 (i.e., completely agree). Their gaze behaviour was recorded as they watched the video clips using an EyeLink 1000 Plus eye-tracker. Dispersion, defined as the mean distance from the gaze coordinate of a given scenario, appeared to be significantly different among the six eHMI conditions. To be more specific, a significantly higher dispersion was observed for the projection condition, and a significantly lower one for the wheels, except when compared to the windscreen. In general, the roof, windscreen, and grill-based eHMIs gave the best performances, except when the car was coming from a corner. The authors therefore recommend the use of omnidirectional external interfaces.

IV. DISCUSSION

In this section, we will first discuss on the visual exploration behaviours of different population groups in street-crossing situations. Then, the methodological approaches adopted in the works presented in this article will be explored. The main eye metrics analysed in these studies to examine both visual exploration behaviour and cognition will be reviewed. Finally, the use of the eye-tracking technology to study interactions between pedestrians and future automated vehicles will be considered.

A. VISUAL SEARCH BEHAVIOURS OF DISTINCT POPULATION GROUPS

Several works studying the visual search strategies of children when compared to that of adults were found in the literature (i.e., Biassoni et al. [22], Egan et al. [23], [24], and Tapiro et al. [25], [27]), as children unfortunately happen to constitute a large group in pedestrian accidents. Findings from these studies consistently showed that the visual exploration of children is more reduced than that of adults. Predominantly, children presented shorter fixation duration and higher fixation frequency than adults. Moreover, they also looked at more objects considered irrelevant to safely cross the road. One possible explanation for these results is that the visual field of children is more reduced than that of adults [33]. They also have less experience, and their...
perceptual and cognitive skills are still being developed. As previously shown in the literature, children are therefore less able to make safe crossing decisions compared to adults, or even teenagers [34]. Some authors (e.g., Biassoni et al. [22], and Tapiro et al. [27]) concluded that training programs, using virtual reality for instance, could help children to develop effective visual exploration strategies.

With population aging and ecological concerns, the number of older pedestrians will increase in the coming years [35]. Although difficulties for older pedestrians to make safe street-crossing decisions were shown in the literature (e.g., [36], [37]), only two studies explored the visual strategies of older adults during the decision-making process of crossing a street using an eye-tracker (i.e., Tapiro et al. [26], and Zito et al. [28]). These eye-tracking studies aimed to expand existing knowledge about decision-making process during street-crossing, and may have practical implications in terms of speed limits, road design, or even pedestrian training programs. Metrics such as gaze distribution, percentage of fixations, as well as the number of head movements were used to explore the visual exploration behaviour of older adults. In general, compared to young adults, the older ones more often looked at the ground than peripheral areas. This could be partly due to cognitive and visual difficulties with increasing age [38], reduced motor abilities such as rigidity in the neck [39], and reduced visual field of view [40]. The authors concluded that visual exploration behaviour associated with other factors may lead older adults to make wrong crossing decisions.

To our best knowledge, no study has investigated visual exploration strategies of pedestrians with disabilities in street-crossing situations using the eye-tracking technology. Notwithstanding, one study explored the visual search strategies of pedestrians with, and without, visual and cognitive impairments in a shared zone [41]. These zones are defined as priority areas for pedestrians, with the goal of encouraging engagement between all road users. Findings showed that individuals with visual and cognitive impairments need more time to process the surrounding environment.

Further studies investigating the visual exploration of different population groups in street-crossing situations need to be conducted in order to provide new information on how individuals process their environment to decide, or not, to cross.

B. METHODOLOGICAL CONSIDERATIONS

Different methodological approaches were employed by the authors, i.e., the use of a real environment (Egan et al. [23], [24], Dey et al. [29], Geruschat et al. [31], and Jiang et al. [32]), of images and videos (Biassoni et al. [22], and Eisma et al. [30]), and of virtual reality (Tapiro et al. [25]–[27], and Zito et al. [28]). Figure 3 represents an example of each type of experimental settings. These situational contexts present benefits and disadvantages. Indeed, experiments carried out on real roads allow a very realistic and naturalistic representation of one’s behaviour. However, such methodology may appear dangerous to experimenters and participants, and may be complicated to replicate (e.g., weather conditions). As for still images, they may not reflect real life where pedestrians have to explore their environment in a dynamic context. Virtual reality (VR) offers major advantages in reproducibility and experimental control (when compared to experiments in a real road environment), and immersivity (unlike images and videos). The main benefit of VR is that it allows users to safely plunge into a simulated, yet realistic, street-crossing situation. However, one should note that VR also presents several limitations, such as its initial costs, potential locomotion sickness, and graphical restrictions which can limit the transferability to real world applications.

It is interesting to note that a few research teams completed their analysis of eye movements and positions by analysing complementary measures. Indeed, in their three studies using the virtual reality technology, Tapiro et al. [25]–[27] asked their participants to press a button when they felt it was safe to cross. In a similar way, Dey et al. [29] made the use of a potentiometer to evaluate, in real time, the participants’ willingness to cross. Eisma et al. [30], on the contrary, administered a questionnaire to their participants following the eye-tracking experiment. Various research methods have been used on their own in the literature to define which factors, and to what extent, affect pedestrians’ decision to cross, or not to cross [42], [43]. However, methods such as observation, interviews, and questionnaires can be affected by the observer bias, or can lead to a subject bias [44]. They yet become extremely interesting when used conjointly with the eye-tracking technology.

It is worth mentioning that no study which combines psychophysics metrics, like speed or distance estimations (e.g., [45], [46]), and eye-tracking measures were found. This kind of studies would likely lead to a better understanding of the relationship between gaze behaviour, and estimation of physical dimensions as made by pedestrians.

C. GAZE ANALYSIS: METRICS AND PROCESSES

In most of the studies presented above, several eye-tracking metrics were used; quantitative metrics related to fixations were the most frequent. Fixations represent the maintaining of the gaze on a single location, for a duration varying between 50 and 600 ms. They are generally assumed to reflect the spatiotemporal location of an object as input to be processed. As the gaze is focused, for a fixation time, onto an object or area in the environment, it is commonly inferred that visual attention is oriented towards it during this time. That way, fixations are usually considered as an indicator of the ongoing attentional processes [22]. More specifically, fixation length may represent the duration of the cognitive processing of the target element, while fixation location is related to the area or the type of information to be processed.

One key eye-tracking analysis procedure was very frequently used: the regions (or areas) of interest (ROIs). Indeed, as street-crossing situations are, most of time, rather complex,
all the elements or areas of a visual scene present different levels of relevance. Dividing the scene into separate ROIs was most of the time necessary. However, as illustrated in Figure 4, the definition of the ROIs was not consistent across articles. While some authors (e.g., Tapiro et al. [25], [26]) defined very global regions (i.e., left, centre, and right), others (e.g., Eisma et al. [30]) were able to define more accurate ROIs (e.g., bumper, hood, and windshield). It is very likely that the accuracy of the definition of these regions was the consequence of several factors, such as: the experimental design, complexity of the infrastructure, or precision of the eye-tracking device chosen.

Gathering location, timing, and duration of fixations allowed the authors to identify specific characteristics of the information intake processes in the studied populations. Some authors used more specific metrics to gain a better understanding of the data. For example, visual attention dispersion, as defined by Tapiro et al. [27], is supposed to reflect the “variety of objects fixated by the pedestrian”. As for the scanning frequency, and the pupil diameter (Jiang et al. [32]), they are assumed to be related to visual alertness and cognitive load, respectively. Finally, transition matrices (Tapiro et al. [26]) were also used to identify visual exploration patterns through the probability of the pedestrian’s gaze to go from one region of interest to another.

The research presented seemed to have mainly focused on where and what is looked at by the pedestrians. The temporal aspect of the visual exploration was less extensively studied. It can be noted that other typical eye-tracking data and tools, such as eye saccades, or saliency maps, have not been used.

D. ON THE INTRODUCTION OF AUTOMATED VEHICLES

The eye-tracking studies presented in this article generally focused on interactions between pedestrians and conventional cars. However, in the near future, the traffic system will be shared between conventional and automated (i.e., without active drivers) vehicles [47]. In order to contribute to a safe environment, one key challenge in the introduction of automated vehicles (AVs) is therefore their interactions with pedestrians. One of the most common causes of accidents involving pedestrians is a misinterpretation of others’ intentions [48]. Several studies have shown the importance of informal (i.e., non-verbal) communication between pedestrians and drivers [18], [49]–[51]. A key concern regarding the introduction of AVs on public roads is therefore due to the changing status of the drivers; AVs may negatively impact interactions with pedestrians, potentially leading to uncertainty and mistrust [52].

To the extent of our knowledge, only very few eye-tracking studies from the pedestrian’s point of view have dealt with the issue of AV. Both Dey et al. [29] and Eisma et al. [30] mentioned the potential benefits of external human-machine interfaces (eHMIs) for the interactions between vehicles and pedestrians, e.g., to inform the latter about the future state of an AV. However, only Eisma et al. carried out an experiment involving such automated vehicles. To summarise, although several articles in the literature investigate the gaze behaviour of AV drivers [21], the pedestrian’s point of view still has to be considerably studied in order to design optimal and effective eHMIs depending on where pedestrians look.

In the near future, using an environment simulated thanks to virtual reality, we will conduct a novel eye-tracking experiment with a view to study the gaze behaviour of both a pedestrian and a driver in the specific context of street-crossing when facing vehicle automation [53].

V. CONCLUSION

In this article, we presented an overview of the existing eye-tracking studies in the literature which investigated the visual exploration strategies of pedestrians in street-crossing situations. More specifically, we analysed the diverse motivations (e.g., investigating the impact of age or of phone use on gaze behaviour), the methodologies employed (e.g., real road environment or virtual reality), and the main findings of the aforesaid studies. In addition, we discussed on several issues, including the population groups involved in the studies, the different research methodologies and gaze metrics used, and the remaining questions raised for the future introduction automated vehicles on our roads. Our survey provides new insights into the visual attention strategies of pedestrians prior to street-crossing, which can be used to inform researchers in traffic psychology.

REFERENCES

[1] A. Borji and L. Itti, “State-of-the-art in visual attention modeling,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 1, pp. 185–207, Jan. 2013.
[2] A. Poole and L. Ball, “Eye tracking in human-computer interaction and usability research: Current status and future prospects,” in Encyclopedia of Human Computer Interaction, Idea Group Reference, Jan. 2006.
[3] L. Lévêque, H. Bosmans, L. Cockmartin, and H. Liu, “State of the art: Eye-tracking studies in medical imaging,” IEEE Access, vol. 6, pp. 37023–37034, 2018.
[4] E. Alemdag and K. Cagiltay, “A systematic review of eye tracking research on multimedia learning,” Comput. Educ., vol. 125, pp. 413–428, Oct. 2018.
[5] J. Gibson, “The theory of affordances,” Perceiving, Acting and Knowing: Toward an Ecological Psychology. Mahwah, NJ, USA: Erlbaum, 1977, pp. 67–82.
[6] J. Gibson, The Ecological Approach to Visual Perception. Psychology Press, 1979, p. 332.
[7] C. Connor, H. Egeith, and S. Yanti, “Visual attention: Bottom-up versus top-down,” Current Biol., vol. 14, pp. 850–852, Oct. 2004.
[8] U. Neisser, Cognition and Reality: Principles and Implications of Cognitive Psychology. New York, NY, USA: Freeman, Dec. 1976, p. 230.
null
LUCIE LÉVÈQUE received the M.Eng. degree in cognitive engineering from the National Superior School of Cognitics, Bordeaux, France, in 2015, the M.Sc. degree in biomedical imaging from the University of Angers, Angers, France, in 2015, and the Ph.D. degree from the School of Computer Science and Engineering, Cardiff University, Cardiff, U.K., in 2019.

She is currently a Research Fellow with the Laboratory of Ergonomics and Cognitive Sciences applied to Transport, Gustave Eiffel University, Lyon, France. Her research interests include visual perception and attention, human–computer interaction, computer vision, and medical imaging. She has been the Leader of Qualinet Task Force 3 (Quality of Experience in medical imaging and healthcare), since 2017.

MAUD RANCHET received the M.Sc. and Ph.D. degrees in cognitive psychology and neuropsychology from Lumière Lyon I University, Lyon, France, in 2008 and 2011, respectively.

She is currently a Research Fellow with the Laboratory of Ergonomics and Cognitive Sciences applied to Transport, Gustave Eiffel University, Lyon, France. Her research interests include cognition and mobility in seniors or adults with neurological disorders (e.g., Parkinson’s disease).

JONATHAN DENIEL received the M.Sc. degree in applied cognitive science from Lumiére Lyon II University, Lyon, France, in 2014, and the Ph.D. degree in cognitive engineering from Bordeaux I University, Bordeaux, in 2019.

He is currently a Research Fellow with the Laboratory of Ergonomics and Cognitive Sciences applied to Transport, Gustave Eiffel University, Lyon. His research interests include human-centred design, advanced driving assistance systems, and driver modeling.

JEAN-CHARLES BORNARD received the M.Eng. degree in cognitive engineering from the National Superior School of Cognitics, Bordeaux, France, in 2009, and the Ph.D. degree in cognitive engineering from Bordeaux I University, Bordeaux, in 2012.

He is currently a Research Engineer with the CIVITEC, Engineering Systems International (ESI) Group, Lyon, France. His research interests include human perception, ergonomics, UX, cognitive modeling, and simulation.

THIERRY BELLET received the M.Sc. degree in cognitive psychology from Lumière Lyon II University, Lyon, France, the M.Eng. degree in artificial intelligence and cognitive sciences from Telecom ParisTech, Paris, France, and the M.Sc. degree in ergonomics and the Ph.D. degree in cognitive ergonomics and cognitive sciences from Paris Descartes University, Paris, in 1998.

He is currently a Senior Researcher at the Laboratory of Ergonomics and Cognitive Sciences applied to Transport, Gustave Eiffel University, Lyon. His research interests include cognitive psychology, cognitive modeling and simulation, driver modeling, human–computer interaction, and artificial intelligence.

Dr. Bellet has been involved in several national and European projects focusing on situational and risk awareness study and modeling including 2BESAFE, PROMOSECUR), human-centred design of adaptive and intelligent co-piloting systems, including CEMVOCAS, ARCOS, COROLA, SAFEMOVE, and ADAS&ME, and cognitive engineering to support adaptive HMI and vehicle automation, including HOLIDES, VI-DAS, and SUAAVE.