Safety Science 181 (2025) 106706

Contents lists available at ScienceDirect

Safety Science

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

Development of a virtual reality pedestrian street-crossing task: The examination of hazard perception and gap acceptance

Joanne M. Bennett [a,b,*,1], Thomas B. McGuckian [a,b,1], Nathan Healy [a], Nikki Lam [a], Ralph Lucas [a], Kathleen Palmer [a], Robert G. Crowther [c], David A. Greene [a], Peter Wilson [a,b], Jonathan Duckworth [d]

[a] School of Behavioural and Health Sciences, Australian Catholic University, NSW, Australia
[b] Healthy Brain and Mind Research Centre, School of Behavioural and Health Sciences, Australian Catholic University, Australia
[c] Exercise and Sport Science, School of Science and Technology, University of New England, Armidale, New South Wales, Australia
[d] School of Design, College of Design and Social Context, RMIT University, Australia

A R T I C L E I N F O

Keywords:
Pedestrian safety
Road crossing
Test development
Vulnerable road user

A B S T R A C T

Pedestrians are a particularly vulnerable road user due to their lack of protection in the event of a crash, which makes safe road-crossing imperative. Current research on pedestrian hazard perception behaviour is limited because street-crossing tasks have not been developed using established procedures. The current study aimed to apply established driver hazard perception principles to the development of a virtual-reality pedestrian street-crossing task (VR-PSCT) which assessed hazard perception and gap acceptance separately. Un-staged street-crossing scenarios (including 36 hazard perception and 41 gap acceptance clips) were filmed at average child and adult heights using 360-degree video cameras at 24 locations across Sydney and Melbourne suburbs. Using established test creation procedures, 16 hazard perception and 17 gap acceptance clips were tested with 76 participants: 32 children ($M = 9.48$, $SD = 1.31$, 75.2 % male) and 44 adults ($M = 23.45$, $SD = 3.46$, 48 % male). Analysis of performance resulted in the removal of another nine clips, resulting in a final VR-PSCT comprising 13 hazard perception and 11 gap acceptance clips. Adults responded more often within the designated hazard and gap windows, had significantly faster response times, and accurately identified hazards more often than children. This indicates that a comparison between adults and children is a useful metric for determining clip inclusion in pedestrian tasks and provides support for the VR-PSCT being an appropriate assessment of two key pedestrian street-crossing behaviours that can be used in future research on pedestrian road safety.

1. Introduction

Pedestrians are individuals who travel on foot by walking, running, or utilising equipment such as canes and walking frames to assist mobility (World Health Organisation [WHO], 2013). Pedestrians are particularly vulnerable to injury as they lack any form of protection in the event of a crash (Unterberger, 2015). Fatalities resulting from road traffic crashes rank as the eighth-leading cause of death globally, with 23 % of these deaths involving pedestrians (WHO, 2020). Pedestrians travel only short distances in comparison to other road users, and their risk of being involved in a crash is only high when they are actively crossing the road (Lassarre et al., 2007; Li et al., 2020). Given this means that pedestrians have a lower risk exposure than other road users, they account for a disproportionate number of fatalities (Office of Road Safety, 2021). While driver behaviour plays a substantial role in pedestrian related crashes, the behaviours of the pedestrian themselves can also be a contributor, with research demonstrating that pedestrians are at-fault for approximately 30% of crashes (Dommes & Cavallo, 2011; Haleem, Alluri & Gan, 2015). It is therefore important to investigate the pedestrian behaviours that may contribute to their involvement in a crash.

1.1. Pedestrian street-crossing behaviours

Pedestrian street-crossing behaviours can be defined as the processes involved in the decision making and transportation between two points

* Corresponding author at: School of Behavioural and Health Sciences, Australian Catholic University, 25A Barker Road, Strathfield, NSW 2135, Australia.
E-mail address: joanne.bennett@acu.edu.au (J.M. Bennett).
[1] Authorship contribution is weighted equally between these two authors as co first authors.

https://doi.org/10.1016/j.ssci.2024.106706
Received 4 March 2024; Received in revised form 17 October 2024; Accepted 21 October 2024
Available online 28 October 2024

(Yıldırım and Celik, 2023). Pedestrian behaviours that underpin this process can be observed during the three stages of a street crossing: (1) Initial Stage – Before Crossing, (2) Middle Stage – Initiation of Crossing, and (3) Final Stage – Completion of Crossing. Various behaviours can be examined to determine pedestrian safety when crossing the road including choice of crossing location, walking speed and distance, and crossing time (Asher et al., 2012; Montufar et al., 2007; Rothman et al., 2012). There are also a number of factors which can influence decision-making for both children and adults when crossing the street including, the distance away the vehicle is (Liu & Tung, 2014), the speed of the approaching vehicle (Dommes & Cavallo, 2011), the number of lanes (Kadali & Vedagiri, 2020), and whether any distractors are present (Leung et al., 2021). This study will focus on key behaviours that occur at the initial and middle stages of crossing. These stages are critical because this is when the individual is standing on the side of the road and making the decision to cross. Pedestrians need to survey the roadway to determine whether it would be safe to initiate crossing, which involves identifying hazards and determining a safe gap between moving vehicles to cross the road. *Hazard perception* (HP) encompasses the identification, interpretation, and subsequent response to a hazard during the initial stage of a road crossing (e.g., approaching vehicles; Rosenbloom et al., 2015). *Gap acceptance* (GA) is the point at which a pedestrian will judge that there is a sufficient gap in the flow of traffic in which they can safely cross the road (Oxley et al., 2005).

HP and GA are related but distinct skills. Although selecting a safe gap to cross requires an awareness of hazards, recognising a hazard does not necessarily dictate a pedestrian's decision to either cross or wait. For example, a pedestrian may see a potential hazard but still opt to cross the road because they have adequate time to cross the road. Conversely, a pedestrian might decide to cross without recognising a hazard. Being able to differentiate between GA and HP is therefore necessary when examining street-crossing safety and the potential predictors of these behaviours.

1.2. Street-crossing tasks

Research investigating street-crossing behaviours has used a range of different methods including questionnaires, static pictures, roadside crossing tasks, computer-based tasks, simulators, virtual reality (VR) and 360-degree video footage (Shen, Ma & Wang, 2023), however, the robustness of some of these methods is questionable. For example, questionnaires and static pictures lack ecological validity (Shen, Ma & Wang, 2023), and computer-based tasks do not elicit representative exploratory behaviours (e.g., head movement; Deb et al., 2017). The use of VR to examine street-crossing behaviours is becoming more common as the technology enables more realistic representation of the complexity of the roadway environment and enables participants the freedom to explore the 360-degree roadway environment (Tapiro, Oron-Gilad & Parmet, 2020). However, computer-generated imagery displayed in a VR headset may not elicit representative thresholds for recognition of danger (Crundall et al., 2021). In addition to computer-generated imagery, VR headsets can display 360-degree video footage, which enables researchers to display real-world visual stimuli (i.e., street-crossing situations) in a safe setting while allowing exploration of the situation that is more natural to human visual and auditory senses. Whilst VR could be beneficial in improving the measurement and therefore our understanding of pedestrian behaviours, there are a lack of studies that provide guidance on best practice principles for the development of VR street-crossing tasks (Shen, Ma & Wang, 2023).

1.3. Comparing Populations to develop street-crossing tasks

The methods for developing and validating HP tasks for drivers has been well documented (Moran et al., 2019). The primary assumption made in the development of HP tests is that younger more inexperienced drivers will perform more poorly on HP tests than older more experienced drivers (Horswill & McKenna, 2004; Wetton, Hill & Horswill, 2011). Research has consistently demonstrated that less experienced drivers have longer response times to hazards than experienced drivers (Horswill, 2016). This difference in response time between these population groups has been the main metric used to determine which scenarios are included in HP tests (Scialfa et al., 2012; Moran et al., 2019), despite the exact participant groups varying in age and experience between studies (Moran et al., 2019). There is currently no established criteria in the pedestrian literature.

Defining inexperience as a pedestrian is more challenging than for drivers because people interact with the roadway on foot from a young age and there is no licensing requirement to gauge experience among pedestrians. Comparative studies have consistently shown that children under the age of 13 years have slower HP response times than middle-aged adults aged 18–54 years (Meyer et al., 2014; Rosenbloom et al., 2015), and the latter out-perform older adults (aged above 65 years; Rosenbloom et al., 2015). A similar pattern has been established when examining pedestrian GA, with children and older adults having poorer GA (i.e., taking longer to select an appropriate gap) than middle-aged adults (Ishaque & Noland, 2008). Following methods used in driving research, comparison between children under the age of 13 years (or inexperienced pedestrians) and adults in middle-adulthood (experienced pedestrians) could index the inclusion of appropriate scenarios that can discriminate performance in a pedestrian street-crossing task.

1.4. Additional principles for test development

Wetton, Hill and Horswill (2011) outline five principles of HP test creation that should be used when developing a pedestrian street-crossing task. The first principle is that HP tests should only measure the construct of HP skill. However, as argued by Horswill and McKenna (2004), HP and GA are separate constructs because GA encompasses risk-taking propensity—as such, they should be measured separately (Horswill & McKenna, 2004). This distinction between HP and GA has been operationalised in a recent driving study by Horswill et al., (2021), whereby separate HP and GA tasks were provided to participants. The HP task required participants to click as soon as possible on any road users who were likely to become involved in a conflict with the car. The GA task required participants to click when they considered there was a sufficient gap and would be willing to turn across traffic. In both tasks response time was measured. To date, a similar approach of having distinct HP and GA tasks has not been used in the pedestrian literature.

The second principle is that tests should discriminate between individuals based on differences in skill and not simply response time ability. This requires that the clips include anticipatory informational cues that experienced road users would identify and use to make decisions earlier than inexperienced road users (Wetton et al., 2011). Clips that include 'pop-out' hazards (e.g., a cyclist obstructed by parked cars suddenly coming into view) would capture simple response time to the hazard but not distinguish those with better HP skill. The third principle is that videos used in the tasks should present un-staged situations. Unstaged videos allow for the presentation of genuine hazardous situations complete with the anticipatory cues that are present in real-life scenarios. The fourth principle is that task instructions need to unambiguously define what situations warrant a response. Wetton et al. (2011) argue that individuals have preconceived ideas about what is a hazardous situation. Wetton et al. (2011) argue that defining a hazard as a *likely* hazard is a useful mechanism to ensure that participants do not wait too long before responding to a genuine hazard. Finally, a HP test needs to be able to identify inappropriate responses to mitigate cheating (Wetton et al., 2011). For button-press style tasks, recommendations have been made to ensure tight response windows are created to minimise the chance of capturing inappropriate responses as genuine (Jackson et al., 2009). However, it is also important to ensure that the window is not so tight that it excludes early responses from experienced road users.

Whilst these principles have been guiding the development of HP tasks in drivers for over a decade, to date these principles have not been applied to pedestrian research. Testing the applicability of these best practice principles in the development of discrete HP and GA tasks for pedestrians would be beneficial for determining their usability in future research.

1.5. Aim

To date, research investigating pedestrian HP has been limited by the absence of the use of established procedures when developing the tasks. The development of the following street-crossing task which separates out HP from GA applied the five principles of HP test creation for drivers and proposes using a comparison between children and adults as a method for determining clip inclusion in the task. Given that children have poorer performance on street-crossing tasks than adults (Meyer et al., 2014; Rosenbloom et al., 2015), we expected children to have a higher proportion of responses outside the response windows, and slower responses times, than adults. The aim of the current study was therefore to apply the best practice principles used for driving HP test development to a VR pedestrian street-crossing task (VR-PSCT) which examines critical pedestrian safety behaviours of HP and GA in children and adults.

2. Method

A methodological diagram depicting the steps taken throughout the development of the VR-PSCT is provided in Fig. 1.

2.1. Stage One: Development of clips

A thorough review of the literature was conducted by four research team members to canvas the types of roadway environments which should be included in the VR-PSCT. This literature review aimed to identify the environments and roadway situations which have been linked to increased crash rates in pedestrians across the lifespan. The research team discussed the findings of the review and compiled a list of roadway environments that are present in Australia that would be included in filming. The list was also presented to an independent expert who works in a government road safety research role for review. The

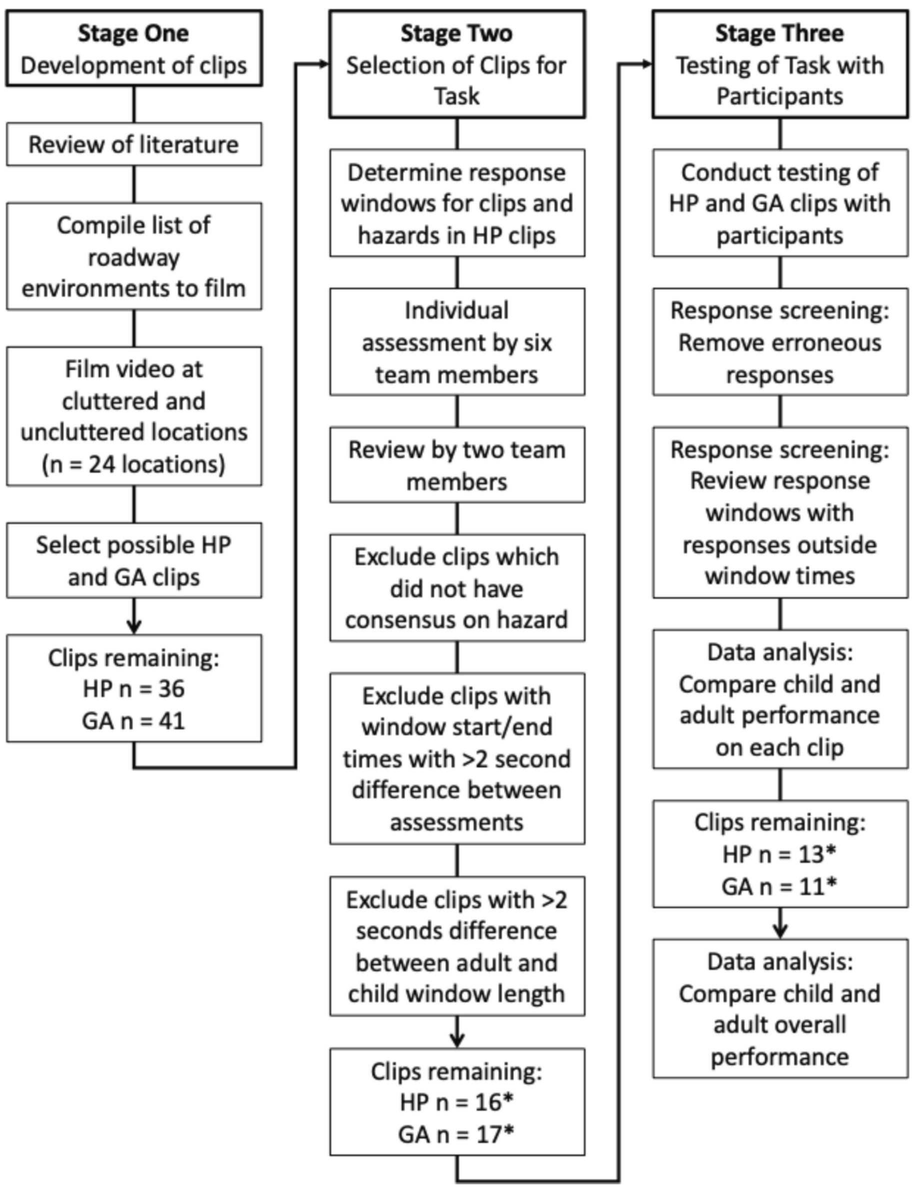

Fig. 1. *Flow chart of methodological steps. Note.* * indicates the number of clips remaining, excluding two practice clips and two clips where no response is expected.

following 12 roadway environments were selected for filming:

- Single-lane two-way (single lane of traffic in each direction)
- Single-lane two-way with bike lane (single lane of traffic and a bicycle lane in each direction)
- Single-lane two-way with refuge (single lane of traffic in each direction, separated by a pedestrian refuge)
- Single-lane two-way with tram (single lane of traffic with a tram line in each direction)
- Single-lane one-way (single lane of traffic going in one direction)
- Single-lane one-way curved (single lane of traffic going in one direction on a curved road)
- Single-lane one-way cycle heavy (single lane of traffic with a bicycle lane going in one direction. Situation has a lot of cyclists using a road)
- Multi-lane two-way (two lanes of traffic in each direction)
- Four-way intersection (single lane of traffic in each direction of four-way unsignalised intersection)
- Pedestrian crossing (single lane of traffic in each direction with pedestrian crossing)
- Roundabout (single lane of traffic in each direction of a four-way roundabout)
- T-intersection (single lane of traffic in each direction of unsignalised t-intersection)

For each of the aforementioned roadway environments, clips were filmed at a location which was considered uncluttered (no obstruction to the view of the roadway) and cluttered (the roadway was partially obscured by parked cars or objects; Peel et al., 2023; Tapiro, Oron-Gilad & Parmet, 2020). This resulted in filming being completed at 24 unique locations. At each location footage was captured for the HP component of the task and the GA component of the task.

To capture un-staged video footage of road scenarios that would provide a full 360-degree view of the roadway environment, INSTA360 X3 (Arashi Vision Inc., Shenzhen, China) 360-degree video cameras were used. Two cameras captured footage simultaneously, one set at the height of the average Australian child (1.32 m) for the child participants and one at the height of the average Australian adult (1.71 m) for the adult participants. The footage was captured from the pedestrian's perspective on the side of the road at various locations across Sydney and Melbourne to provide a variety of road environments (e.g., trams more commonly operate in Melbourne than Sydney) and ensure increased ecological validity for pedestrians across Australia. This filming was always completed during daylight hours with consistent weather conditions (i.e., full sunlight with no rain) across all filming locations. Filming was completed by teams of three research team members at all sites. Team members individually recorded the time of filming where they identified a possible HP or GA scenario. Multiple potential scenarios at each roadway environment were identified. Some locations resulted in both HP and GA scenarios being captured at the same location.

The footage from each roadway environment was rewatched by two research team members and possible HP or GA scenarios were selected for editing. This resulted in 36 possible HP scenarios and 41 possible GA scenarios being created. The same scenarios were used for both the child and adult videos. These videos were edited into shorter clips of varying lengths using Adobe Premiere Pro (Adobe Inc., San Jose, USA) and exported as 5760x2880 H.264 video files at 30fps using Adobe Media Encoder (Adobe Inc., San Jose, USA).

2.2. Stage Two: Selection of clips for task

Every clip was reviewed by three research team members to select appropriate clips and to determine the associated response windows for each clip. Three team members watched the adult clips, and three different team members watched the child clips. Each researcher watched the clips independently and were required to identify the hazard or gap window start and finish in seconds from the start of the clip. For HP clips, researchers also identified the hazard present in the clip (e.g., white car approaching from right). The HP window start times were defined as the initial emergence of a potential hazard, and the HP window end was defined as the point in which the hazard was directly in line with the path that the participant would have taken to cross the road. The GA window start times were defined as the initial emergence of a safe gap to cross the road, and the GA window end time was defined as the point in which the vehicle was directly in line with the path that the participant would have taken to cross the road.

Two research team members met and reviewed the clips and window times. Firstly, in line with principle four (Wetton et al., 2011), clips that did not have consensus on the potential hazard were excluded due to ambiguity in the scenario. Secondly, window times were reviewed and clips with large discrepancies (a difference of more than 2 s between reviewers) in the window start or end time were excluded. Finally, clips which had large discrepancies in window times (that had a difference of more than 2 s in length of window) between the children and adult clips were excluded. For example, a clip whereby the hazard was visible much earlier in the adult clip (due to not being impacted by roadway clutter given the height of the camera) was excluded as it was deemed to be a much easier task for adults than for children. Only one HP or GA clip was selected from a single location, however the same location may have resulted in both a HP and GA clip.

The above steps resulted in the exclusion of 16 HP clips, resulting in 20 clips for testing in the HP task: two practice clips, 16 clips with a hazard, and two clips that did not contain a hazard to assess acquiescent responding. A description of the 16 clips with a hazard is presented in Table 1. A visual depiction of an example HP scenario is presented in

Table 1
Description and Hazard Response Windows of the Hazard Perception Test Clips.

Clip	Duration (s)	Roadway Environment	Hazardous Event Description	Hazard Window (s)
1	24	Single-lane two-way (cluttered)	Motorcycle	13–23
2	23	Single-lane two-way (uncluttered)	Silver car	8–18
3	24	Single-lane one-way cycle heavy (uncluttered)	Cyclist	12–23
4	12	Four-way intersection (cluttered)	White car	5–11
5	22	Four-way intersection (uncluttered)	Ute	5–18
6	23	Multi-lane two-way (cluttered)	White truck	3–18
7	17	Single-lane one-way (cluttered)	Parked car leaving car spot	6–15
8	15	Single-lane one-way (uncluttered)	White/grey car	5–14
9	27	Pedestrian crossing (uncluttered)	White car	0–17
10	24	Roundabout (uncluttered)	White van	11–23
11	29	T-intersection (cluttered)	Grey/black car	8–18
12	9	T-intersection (uncluttered)	Red car	4–7
13	28	Single-lane two-way with bike lane (uncluttered)	Red car	11–22
14	38	Single-lane two-way with refuge (uncluttered)	Grey/black car	13–29
15	25	Single-lane two-way with tram (cluttered)	White car	8–15
16	30	Single-lane two-way with tram (uncluttered)	Tram	6–22

Note. s = seconds.

Fig. 2. *Visual Depiction of Hazard Perception Clip Response Window. Note.* Screenshots from HP clip number 13. From top to bottom, initially there is no hazard present. A red car (circled) enters the road, this is the identified hazard and signifies the start of the hazard response window. The point in which the hazard would have struck the pedestrian if they were on the road is the end of the hazard window. The image is slightly distorted due to the nature of taking the image from 360-degree footage to a flat two-dimensional (2D) image. Videos are not distorted when viewed in the VR headset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 2. Twenty GA clips were excluded, resulting in 21 clips for testing in the GA task: two practice clips, 17 clips featuring a gap (Table 2), and two clips showcasing a continuous flow of traffic (i.e., no safe crossing gap) to assess acquiescent responding. A visual depiction of a GA scenario is presented in Fig. 3.

2.3. *Stage Three: Testing of task with participants*

2.3.1. *Participants*

Participants in this study consisted of 32 children, ranging in age from 7 to 11 years ($M = 9.48$, $SD = 1.31$, 75.2 % male), and 44 adults ranging in age from 18 to 36 years ($M = 23.45$, $SD = 3.46$, 48 % male). Participants were recruited from the community in both Melbourne and Sydney through a combination of word of mouth and snowball sampling. All participants completed a single testing session in a laboratory located on either the Sydney or Melbourne campus of the Australian Catholic University. 32 participants were tested in Sydney (19 adults, 13 children) and 44 were tested in Melbourne (25 adults, 19 children). Participants were excluded if they had a diagnosis of epilepsy and were required to have English proficiency. All participants were offered a $15 shopping voucher as compensation. No participants exhibited signs of simulation sicknesses and therefore all participants completed the study session. Australian Catholic University Human Research Ethics Committee approved this study (2023-3050H).

Table 2
Gap Acceptance Windows of the Gap Acceptance Test Clips.

Clip Number	Duration (s)	Roadway Environment	Gap Window (s)
1	18	Single-lane two-way (cluttered)	10–16
2	26	Single-lane two-way (uncluttered)	11–22
3	22	Single-lane one-way curved (cluttered)	9–16
4	24	Single-lane one-way cycle heavy (uncluttered)	12–17
5	23	Four-way intersection (cluttered)	18–23
6	23	Four-way intersection (uncluttered)	19–23
7	31	Multi-lane two-way (uncluttered)	0–12
8	31	Multi-lane two-way (cluttered)	15–26
9	24	Single-lane one-way (uncluttered)	7–24
10	23	Pedestrian crossing (cluttered)	5–20
11	23	Roundabout (cluttered)	10–16
12	33	Roundabout (uncluttered)	24–33
13	25	T-intersection (cluttered)	7–16
14	28	Single-lane two-way with bike lane (uncluttered)	13–21
15	22	Single-lane two-way with refuge (cluttered)	11–16
16	16	Single-lane two-way with refuge (uncluttered)	6–16
17	25	Single-lane two-way with tram (cluttered)	6–17

Note. s = seconds.

2.3.2. *VR-PSCT*

Participants completed the pedestrian street crossing task using a head-mounted VivePro Eye headset (HTC, Taoyuan City, Taiwan) powered by an Alienware X17 laptop (Dell Technologies, Round Rock, USA). The VivePro Eye displayed the 360-degree video clips via a custom-built Unity project (Unity Technologies, San Francisco, USA). In line with principle one (Wetton et al., 2011), two versions of the street-crossing task were administered: the HP task and the GA task. Presentation of these tasks was counterbalanced across participants to prevent order effects. Response time was the primary outcome used to measure performance in both tasks. The time at which the participant pulled the controller trigger was recorded by the Unity project. Response time (in seconds) was calculated as the time between the window start and the point at which the participants pulled the controller trigger. Faster response times indicated better HP and GA skill.

2.3.2.1. *Hazard perception task.* Participants were instructed to imagine they were the pedestrian on the side of the road waiting to cross to the other side of the street. They were informed that they could look around the entire area freely. They were instructed to pull on the controller trigger when they first identified a potential hazard on the road which would make it unsafe to cross the street. A hazard was defined as either an object or event present in the road environment which would make it unsafe to cross the street safely. Following every clip, participants were asked to name the hazard they responded to. Responses were marked as correct or incorrect with higher scores indicating better hazard accuracy.

2.3.2.2. *Gap acceptance task.* Participants were instructed to imagine they were the pedestrian on the side of the road waiting to cross to the other side of the street. They were informed that they could look around the entire area freely. They were instructed to pull the controller trigger to identify when it would be safe to cross the road safely at their normal walking pace (i.e., if they would have to run at any time to make it across safely, then that would not be a safe crossing time).

2.3.3. *Procedure*

Participants were assessed individually in face-to-face sessions. Upon

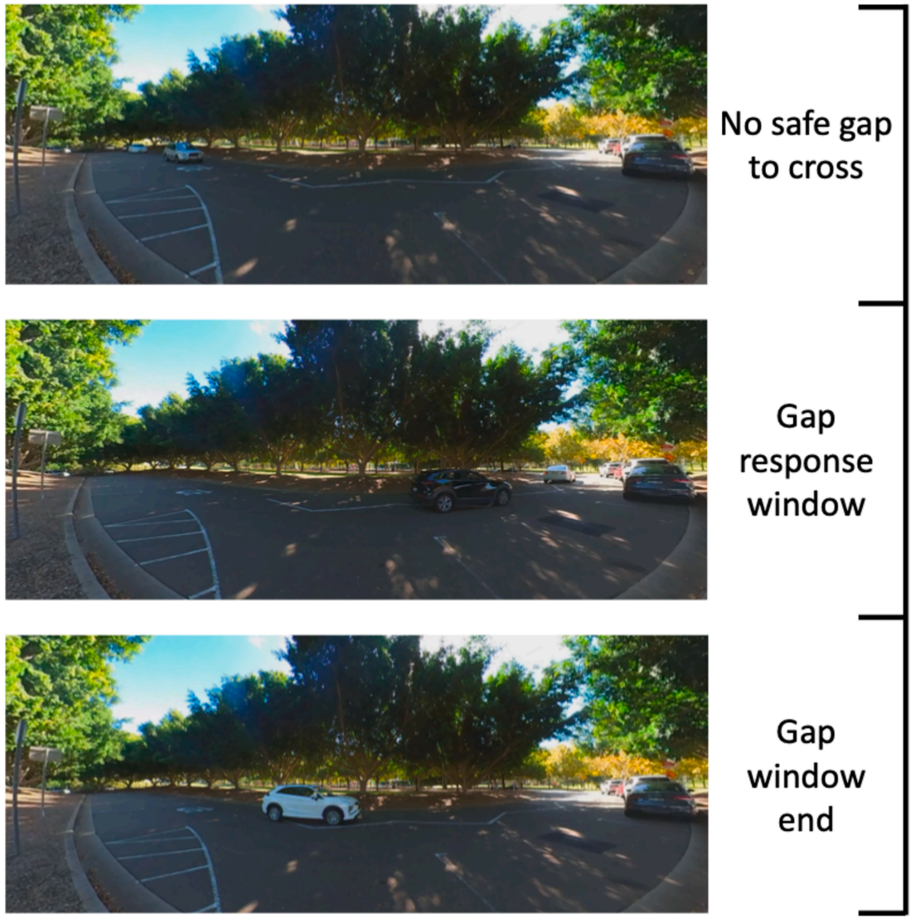

Fig. 3. Visual Depiction of Gap Acceptance Clip Response Window. *Note.* Screenshots from GA clip number 9. From top to bottom, initially there is no gap in the traffic. A gap in the traffic then appears and signifies the start of the gap acceptance window. The point at which the car is directly in line with the path that the participant would have crossed the road is the end of the gap acceptance window. The image is slightly distorted due to the nature of taking the image from 360-degree footage to a flat 2D image. Videos are not distorted when viewed in the VR headset.

arrival, adult participants provided informed consent to participation, and for the children, a parent or guardian firstly provide consent and children were also required to complete an assent form. Following this all participants completed a demographic questionnaire, with children able to complete this with their parent or guardian. Participants completed the pedestrian street crossing task using the VR headset. First the headset was calibrated to ensure comfort and clarity of the visual display for each participant. The order of the street crossing tasks was counterbalanced. Before each of the two tasks participants completed practice clips where they were given feedback on the required actions to be taken during the clip and an opportunity to ask questions. This enabled researchers to check that participants understood the requirements of the task. Participants were offered a break between each street crossing task and were monitored for simulation sickness using a simulation sickness questionnaire (Kennedy et al., 1993). The VR testing session lasted approximately 30 mins.

2.3.4. *Data screening*

The stages of the data analysis were in keeping with the HP principles outlined by Wetton et al., (2011) and followed previous HP driving research (Moran et al., 2019). This was done in two stages; firstly data screening of the responses to individual clips and secondly data analysis comparing the performance of children to adults per clip.

2.3.4.1. Response screening method. To conduct the screening of responses to individual clips we firstly identify the potential for multiple hazards/gaps being present in the clips. This was done through an analysis of the percentage of participants responding more than once for each clip. To examine erroneous responding we identified participants that responded more than 10 times within a single clip. These participants' responses were removed from subsequent analysis as the high rate of responding potentially indicated a lack of understanding of the task requirements. To identify whether the windows were appropriate for each of the scenarios an analysis of the pattern of responses relative to the response windows (before, during, and after) was conducted. For this analysis we only used participants' first response for each clip. In line with prior HP research (Moran et al., 2019) where more than two participants responded before the beginning of the window, the clip was reviewed to determine if post-hoc adjustments needed to be made to the windows (i.e., the window time widen).

2.3.4.2. Response screening results. An investigation of the pattern of responses of participants within each clip was conducted for both the HP clips (see Table 3) and the GA clips (see Table 4). Children made multiple responses at a higher frequency than adults. The median number of responses was two, with further analysis demonstrating that these responses often occurred in quick succession. This suggests that multiple responses are likely the result of participants ensuring that they recorded a response rather than being evidence of multiple hazards/gaps.

Children were found to have a higher proportion of responses outside the windows than adults for all clips in both HP and GA tasks. Therefore, response window review was based on more than two adults (i.e., $\geq$ 9.1 %) responding either before or after the response window. Five HP clips (clip 3, 4, 10, 14, and 16) met this criterion for responses before the

Table 3
Analysis of Number of Multiple Responses, Responses Relative to Initial Windows, and Post-hoc window times for Hazard Perception Clips.

Clip Number	Roadway Environment	Group	% Multiple Responses	% Greater than 10 Responses	% Respond Before Window	Mean (SD) time before window (s)	% Respond During Window	% Respond After Window	Mean (SD) time after window (s)	% No Response	Post-hoc window[a]
1	Single-lane two-way (cluttered)	Child	12.5	3.1	6.7	6.73 (8.84)	66.7	20.0	0.62 (0.45)	6.7	13–23
		Adult	15.9	0	0	–	95.5	4.5	0.21 (0.25)	0	
2	Single-lane two-way (uncluttered)	Child	71.8	9.4	20	1.15 (0.90)	66.7	10.0	0.40 (0.19)	3.3	8–18
		Adult	45.5	0	0	–	97.7	2.3	0.24	0	
3	Single-lane one-way cycle heavy (uncluttered)	Child	53.1	3.1	16.7	6.04 (3.50)	70.0	10.0	0.61 (0.34)	3.3	**9–23**
		Adult	34.1	0	25	3.39 (2.49)	70.5	4.5	0.51 (0.44)	0	
4	Four-way intersection (cluttered)	Child	43.8	6.3	0	–	76.7	20.0	0.75 (0.55)	3.3	**4–11**
		Adult	22.7	0	9.1	0.34 (0.24)	86.4	4.5	0.87 (0.31)	0	
5	Four-way intersection (uncluttered)	Child	65.6	6.3	6.7	0.65 (0.06)	90.0	0	–	3.3	5–18
		Adult	25.0	0	2.3	1.63	95.3	2.3	0.29	0	
6	Multi-lane two-way (cluttered)	Child	62.5	6.3	6.7	2.96 (0.04)	80.0	10.0	1.63 (0.95)	3.3	3–18
		Adult	38.6	0	0	–	100	0	–	0	
7	Single-lane one-way (cluttered)	Child	28.1	3.1	3.3	1.21	76.7	13.3	0.65 (0.37)	6.7	6–15
		Adult	2.3	0	0	–	97.7	2.3	0.21	0	
8	Single-lane one-way (uncluttered)	Child	25.0	6.3	6.7	4.99 (0.001)	83.3	6.7	0.54 (0.45)	0	5–14
		Adult	0	0	0	–	97.7	2.3	0.60	0	
9	Pedestrian crossing (uncluttered)	Child	50.0	9.4	0	–	90.0	3.3	1.19	6.7	
		Adult	13.6	0	0	–	97.7	2.3	0.82	0	0–17
10	Roundabout (uncluttered)	Child	56.3	3.1	23.3	1.73 (1.85)	60.0	6.7	0.12 (0.04)	0	**10–23**
		Adult	22.7	0	9.1	0.54 (0.35)	88.6	2.3	0.18	0	
11	T-intersection (cluttered)	Child	68.8	9.4	16.7	4.76 (2.16)	66.7	10.0	1.27 (1.27)	6.7	8–18
		Adult	22.7	0	0	–	95.5	4.5	1.42 (1.90)	0	
12	T-intersection (uncluttered)	Child	21.9	9.4	6.7	2.12 (1.49)	63.3	26.7	0.77 (0.58)	3.3	4–7
		Adult	4.5	0	2.3	2.99	81.8	15.9	1.55 (0.90)	0	
13	Single-lane two-way with bike lane (uncluttered)	Child	68.8	12.5	13.3	4.68 (4.48)	76.7	6.7	2.03 (2.40)	3.3	11–22
		Adult	36.4	0	2.3	3.10	95.3	2.3	0.25	0	
14	Single-lane two-way with refuge (uncluttered)	Child	68.8	15.6	10	3.82 (2.56)	76.7	10	0.40 (0.30)	3.3	**10–29**
		Adult	25.0	0	9.1	3.03 (2.14)	90.9	0	–	0	
15	Single-lane two-way with tram (cluttered)	Child	75.0	18.75	33.3	3.86 (2.96)	56.7	6.7	0.88 (0.25)	3.3	8–15
		Adult	25.0	0	0	–	97.7	2.3	1.39	0	
16	Single-lane two-way with tram (uncluttered)	Child	63.5	15.6	6.7	1.72 (2.40)	83.3	6.7	0.79 (0.63)	3.3	6–22
		Adult	25.0	0	9.1	1.24 (0.83)	88.6	2.3	1.32	0	

Note. [a]Windows that changed post-hoc are in bold, s = seconds.

response window had started. The two research team members examined the mean response time to determine whether a post-hoc change to the window start time needed to occur. For clip 3, 4, 10, and 14 an adjustment was made to the window start time to be inclusive of these responses as the hazard was found to be visible in the far distance. These hazards were not perceptible to the research team when watching the video clips on a flat computer screen, however are visible in the distance when the videos are played through the higher fidelity VR headset. We hypothesised that the mode of presentation of the video clips explains this need for post-hoc adjustment of the window start time. No change was made to clip 16 as a review of the clip found that there was no hazard present at the time participants were responding. One HP clip (clip 12) had 15.9 % of adults responding after the window end time. A review of this clip found that the hazard was clearly opposite the pedestrians at the time they were responding. We hypothesised that due to the nature of the t-intersection that these participants were looking down the wrong street and failed to notice the hazard until it was right opposite them. Therefore, no change was made to the window.

Inspection of GA clips required an assessment about how long a gap needed to be for it to provide sufficient time for safe crossing. The average lane of a street is 3.5 m wide (New South Wales Government, 2023). Using a conservative walking speed of 0.81 m/s (adults aged > 65 years in winter; Montufar et al., 2007), a gap of 4.32 s would be needed to cross a single lane. To allow for an additional safety margin, it was determined that a minimum gap of 5 s would be used as a safe gap window.

For the GA task eight clips met the criterion for further inspection. Two clips (clip 6 and 12) were identified due to responses before the window and six (clip 1, 3, 4, 7, 13 and 17) were identified due to responses after the window. For both clips where participants were responding prior to the window, participants identified small unsafe gaps which would not enable sufficient crossing time (i.e., < 5 s). Therefore, no change was made to the windows for these clips. For five of the six clips with responses after the response window (clip 1, 3, 4, 13 and 17) it appears that participants were waiting until there was no traffic present before choosing to cross. A potential explanation for this is that participants might have been too conservative and therefore missed acceptable safe gaps. In all cases the gap identified by the research team would have enabled sufficient time for a pedestrian to cross the road (i.e., ≥ 5 s). As such, no changes were made to the clip

Table 4
Analysis of Number of Multiple Responses, Responses Relative to Initial Windows, and Post-hoc window times for Gap Acceptance Clips.

Clip Number	Roadway Environment	Group	% Multiple Responses	% Greater than 10 Responses	% Respond Before Window	Mean (SD) time before window (s)	% Respond During Window	% Respond After Window	Mean (SD) time after window (s)	% No Response	Post-hoc window[a]
1	Single-lane two-way (cluttered)	Child	12.5	0	9.4	3.48 (2.84)	18.8	31.3	0.95 (0.38)	40.6	10–16
		Adult	11.4	0	0	–	56.8	20.5	2.46 (0.23)	22.7	
2	Single-lane two-way (uncluttered)	Child	46.9	3.1	6.3	2.88 (3.53)	84.4	9.4	1.73 (0.26)	0	11–22
		Adult	20.5	0	0	–	100	0	–	0	
3	Single-lane one-way curved (cluttered)	Child	37.5	0	15.6	3.89 (1.98)	40.6	28.1	5.26 (2.13)	15.6	9–16
		Adult	20.5	0	2.3	4.33	86.4	11.4	4.77 (0.52)	0	
4	Single-lane one-way cycle heavy (uncluttered)	Child	34.4	6.3	22.6	5.62 (5.21)	29.0	25.8	2.74 (0.51)	22.6	12–17
		Adult	13.6	0	2.3	2.83	20.5	61.4	4.33 (0.79)	15.9	
5	Four-way intersection (cluttered)	Child	43.8	0	25.8	7.35 (6.73)	67.7	0	–	6.5	18–23
		Adult	2.3	0	2.3	0.03	95.5	0	–	2.3	
6	Four-way intersection (uncluttered)	Child	37.5	0	35.5	9.95 (5.12)	54.8	0	–	9.7	19–23
		Adult	11.4	0	15.9	10.94 (1.23)	81.8	0	–	2.3	
7	Multi-lane two-way (uncluttered)	Child	34.4	0	0	–	16.1	45.2	9.40 (6.03)	38.7	0–12
		Adult	9.1	0	0	–	18.2	50	14.07 (4.60)	31.8	
8	Multi-lane two-way (cluttered)	Child	53.1	0	45.2	4.34 (2.77)	51.6	0	–	3.2	15–26
		Adult	15.9	0	0	–	100	0	–	0	
9	Single-lane one-way (uncluttered)	Child	43.8	0	3.3	0.45	96.7	0	–	0	7–24
		Adult	11.4	0	2.3	0.18	97.7	0	–	0	
10	Pedestrian crossing (cluttered)	Child	28.1	3.1	6.7	1.77 (2.01)	73.3	6.7	0.93 (0.79)	13.3	5–20
		Adult	4.5	0	0	–	100	0	–	0	
11	Roundabout (cluttered)	Child	25.0	0	9.7	4.81 (3.64)	29.0	9.7	1.92 (1.08)	51.6	10–16
		Adult	2.3	0	0	–	41.9	4.7	3.30	53.5	
12	Roundabout (uncluttered)	Child	59.4	6.3	87.1	11.77 (4.35)	12.9	0	–	0	24–33
		Adult	15.9	0	77.3	9.98 (5.66)	22.7	0	–	0	
13	T-intersection (cluttered)	Child	53.1	6.3	16.1	1.70 (1.67)	67.7	16.1	3.08 (1.50)	0	7–16
		Adult	22.7	0	4.5	2.31 (0.07)	84.1	11.4	4.74 (0.81)	0	
14	Single-lane two-way with bike lane (uncluttered)	Child	37.5	3.1	19.4	7.60 (3.81)	61.3	9.7	4.74 (1.02)	9.7	13–21
		Adult	4.5	0	0	–	65.9	4.5	5.10 (0.08)	29.5	
15	Single-lane two-way with refuge (cluttered)	Child	21.9	3.1	22.6	3.51 (1.87)	16.1	16.1	2.78 (1.53)	45.2	11–16
		Adult	0	0	4.5	2.18 (0.40)	34.1	0	–	61.4	
16	Single-lane two-way with refuge (uncluttered)	Child	46.9	6.3	3.2	0.77	87.1	0	–	9.7	6–16
		Adult	15.9	0	0	–	100	0	–	0	
17	Single-lane two-way with tram (cluttered)	Child	46.9	0	3.2	1.76	87.1	3.2	3.82	6.5	6–17
		Adult	15.9	0	2.3	0.20	81.4	9.3	3.67 (1.70)	7.0	

Note. [a]No windows were changed post-hoc, s = seconds.

windows. For clip 7 it was found that there was a vehicle which turns down a side street in the distance during the gap window, and it is hypothesised that this might have prevented participants responding during the window. This creates some confusion in the clip which could not be rectified by adjusting the windows and therefore this clip was excluded from further analysis. No changes were made to any GA windows post-hoc.

2.3.5. Data analysis

A comparison of performance for each clip between children and adults was conducted. HP and GA response times that fell within the adjusted windows (Table 3 and Table 4) were included. No response and responses outside response windows were deemed to be legitimate attempts at completing the task and therefore, in line with procedures used in driving HP studies (Moran et al., 2020), scores for these participants were imputed with the maximum acceptable response time (i.e., window end time). Independent samples t-tests (two tailed) were run to compare the response times between children and adults for each clip. Alpha was set at 0.05 for all analyses, with adjustments made for violations of homogeneity of variance where appropriate. Cohen's d effect sizes have been reported for each clip, with 0.2 indicating a small effect, 0.5 a medium effect, and 0.8 a large effect.

3. Results

3.1. Comparison of performance between children and adults

For the HP clips, adults performed significantly faster than children on 13 of the 16 clips, with effect sizes ranging from medium to large (Table 5). Three clips (11, 12, and 15) found no significant difference in performance between children and adults, and two of these clips (12 and 15) the effect was in the direction opposite to expectation with children performing faster than adults. For the GA clips, adults performed significantly faster than children on 11 of the 16 clips, with effect sizes ranging from small to large (Table 6). Five clips (11, 12, 13, 14, and 17) found no significant difference in performance between children and adults.

3.2. Final VR-PSCT

The clips included in the final HP and GA tasks are presented in Table 7. To examine the internal consistency of the two tasks, reliability analyses were conducted on the clips included in the final tasks. For GA tasks, internal consistency was acceptable (Cronbach's $\alpha = 0.77$) and for HP tasks internal consistency was excellent (Cronbach's $\alpha = 0.91$).

To examine performance on the final HP and GA tasks, an overall score for each group was derived for each task using the response times for each clip. The following steps were completed in line with established procedures for calculating total scores in the driving HP literature (Hill et al., 2019; Moran et al., 2020). Firstly, due to differing clip and window lengths for each clip, each participant's response time for each individual clip were standardised as z-scores. Secondly, response time z-scores were averaged across HP clips and GA clips separately. Finally, the averaged response time z-score was then converted back to a raw score using the population mean and standard deviation. To complete the comparative analysis, the average HP response time, and GA response time were calculated for child and adult groups. Furthermore, HP accuracy was obtained per participant by summing the number of hazards correctly identified during the HP clips. Medians for HP accuracy scores were calculated for both the child and adult group. The descriptive statistics for overall HP and GA response time and HP accuracy for adult and child groups are reported in Table 8. Independent t-tests indicate that children have significantly slower overall response times compared to adults for both the HP and GA tasks. A Mann-Whitney *U* test indicated that children were significantly less accurate on the HP task than adults (Table 8).

Pearson correlations were run between the HP and GA overall response time for both children and adults to examine the relationship between the two tasks. Non-significant weak positive relationships were found between HP and GA response times for both children ($r = 0.13$, $p = 0.489$) and adults ($r = 0.26$, $p = 0.139$), suggesting that HP and GA are distinct pedestrian skills.

4. Discussion

4.1. Overview of study

Reductions in pedestrian fatalities have stagnated despite an overall reduction in road trauma, making it more important than ever to understand the behaviours of pedestrians themselves (Levulytė et al., 2017; SWOV, 2020). Whilst a variety of methodologies have been used to measure pedestrian behaviour, there are no established procedures for developing these tasks. This study aimed to apply best practice

Table 5
Hazard Perception Task: Missing Data, Descriptives, and T-test comparison of Children and Adults.

Clip Number	Roadway Environment	Group	% Missing	*M* (SD) Response Time (s)[a]	t-test[b]	Cohen's d
1	Single-lane two-way (cluttered)	Child	40.6	7.90 (2.65)	$t(71) = 2.93$, $p = 0.002$	0.70
		Adult	4.5	6.20 (2.28)		
2	Single-lane two-way (uncluttered)	Child	46.9	10.01 (3.89)	$t(36.3) = 1.86$, $p = 0.036$	0.52
		Adult	2.3	4.98 (2.05)		
3	Single-lane one-way cycle heavy (uncluttered)	Child	31.3	10.01 (3.89)	$t(71) = 2.84$, $p = 0.003$	0.68
		Adult	20.5	7.28 (4.11)		
4	Four-way intersection (cluttered)	Child	34.4	4.87 (1.90)	$t(70) = 3.39$, $p < 0.001$	0.82
		Adult	4.5	3.44 (1.63)		
5	Four-way intersection (uncluttered)	Child	21.9	7.42 (4.02)	$t(69) = 2.26$, $p = 0.013$	0.55
		Adult	6.8	5.46 (3.23)		
6	Multi-lane two-way (cluttered)	Child	28.1	8.76 (5.23)	$t(39.7) = 2.79$, $p = 0.004$	0.75
		Adult	0.0	5.70 (3.17)		
7	Single-lane one-way (cluttered)	Child	31.3	6.62 (1.95)	$t(71) = 2.41$, $p = 0.009$	0.58
		Adult	2.3	5.60 (1.62)		
8	Single-lane one-way (uncluttered)	Child	25.0	5.42 (2.76)	$t(44.6) = 3.98$, $p < 0.001$	1.04
		Adult	2.3	3.02 (1.98)		
9	Pedestrian crossing (uncluttered)	Child	25.0	9.82 (4.49)	$t(43.8) = 2.17$, $p = 0.018$	0.57
		Adult	2.3	7.65 (3.36)		
10	Roundabout (uncluttered)	Child	37.5	8.62 (4.66)	$t(46.1) = 3.11$, $p = 0.002$	0.80
		Adult	2.3	5.53 (3.28)		
11	**T-intersection (cluttered)**	Child	40.6	5.90 (3.63)	**$t(40.9) = 1.07$, $p = 0.147$**	0.29
		Adult	4.5	5.05 (2.48)		
12	**T-intersection (uncluttered)**	Child	50.0	2.34 (0.91)	**$t(69) = -0.92$, $p = 0.181$**	−0.24
		Adult	18.2	2.53 (0.72)		
13	Single-lane two-way with bike lane (uncluttered)	Child	34.4	5.82 (3.48)	$t(40.7) = 1.89$, $p = 0.003$	0.51
		Adult	6.8	4.34 (2.53)		
14	Single-lane two-way with refuge (uncluttered)	Child	34.4	12.14 (5.05)	$t(67) = 2.40$, $p = 0.010$	0.60
		Adult	4.5	9.47 (4.06)		
15	**Single-lane two-way with tram (cluttered)**	Child	50.0	4.85 (2.39)	**$t(69) = -0.74$, $p = 0.230$**	−0.18
		Adult	2.3	5.25 (2.15)		
16	Single-lane two-way with tram (uncluttered)	Child	31.3	8.20 (5.88)	$t(44.21) = 1.97$, $p = 0.027$	0.52
		Adult	11.4	5.60 (4.45)		

Note. Sample size varies per analysis due to the exclusion of participants who demonstrated erroneous responding. Clips with a non-significant difference between children and adults are presented in bold.

[a] Values represent seconds from the beginning of the window.

[b] Corrections for violations of homogeneity of variance have been made where appropriate. s = seconds.

Table 6
Gap Acceptance Task: Missing Data, Descriptives, and T-test comparison between Children and Adults.

Clip Number	Roadway Environment	Group	% Missing	*M* (SD) Response Time (s)[a]	t-test[b]	Cohen's d
1	Single-lane two-way (cluttered)	Child	81.3	5.43 (1.27)	$t(69.9) = 4.72, p < 0.001$	1.01
		Adult	43.2	3.49 (2.27)		
2	Single-lane two-way (uncluttered)	Child	18.8	3.80 (3.45)	$t(32.1) = 3.47, p = 0.002$	0.95
		Adult	0.0	1.62 (0.77)		
3	Single-lane one-way curved (cluttered)	Child	59.4	5.30 (2.19)	$t(59.7) = 4.56, p < 0.001$	1.09
		Adult	13.6	3.13 (1.84)		
4	Single-lane one-way cycle heavy (uncluttered)	Child	75.0	4.13 (1.49)	$t(70.6) = 3.49, p < 0.001$	0.79
		Adult	38.6	2.75 (1.91)		
5	Four-way intersection (cluttered)	Child	34.4	2.78 (1.83)	$t(44.1) = 3.29, p = 0.002$	0.84
		Adult	4.5	1.58 (1.06)		
6	Four-way intersection (uncluttered)	Child	46.9	2.92 (1.12)	$t(73) = 2.85, p = 0.006$	0.67
		Adult	18.2	2.19 (1.07)		
8	Multi-lane two-way (cluttered)	Child	50.0	6.74 (2.49)	$t(51.7) = 6.04, p < 0.001$	1.49
		Adult	0.0	3.57 (1.82)		
9	Single-lane one-way (uncluttered)	Child	9.4	3.63 (3.33)	$t(72) = 2.24, p = 0.028$	0.53
		Adult	2.3	2.11 (2.52)		
10	Pedestrian crossing (cluttered)	Child	34.4	4.95 (2.25)	$t(38.1) = 3.86, p < 0.001$	1.04
		Adult	0.0	3.20 (1.17)		
11	**Roundabout (cluttered)**	Child	71.9	4.72 (2.05)	$\mathbf{t(67.6) = 1.12, p = 0.266}$	0.25
		Adult	59.1	4.16 (2.21)		
12	**Roundabout (uncluttered)**	Child	87.5	8.42 (1.51)	$\mathbf{t(70.8) = 1.46, p = 0.148}$	0.32
		Adult	77.3	7.75 (2.43)		
13	**T-intersection (cluttered)**	Child	37.5	4.94 (3.11)	$\mathbf{t(51.8) = 1.56, p = 0.062}$	0.39
		Adult	15.9	3.85 (2.55)		
14	**Single-lane two-way with bike lane (uncluttered)**	Child	43.8	5.20 (2.63)	$\mathbf{t(72) = 1.33, p = 0.184}$	0.31
		Adult	34.1	4.34 (2.81)		
15	Single-lane two-way with refuge (cluttered)	Child	84.4	4.49 (1.22)	$t(71.2) = 1.80, p = 0.038$	0.40
		Adult	65.9	3.89 (1.62)		
16	Single-lane two-way with refuge (uncluttered)	Child	18.8	4.75 (2.88)	$t(39.2) = 2.47, p = 0.018$	0.65
		Adult	0.0	3.35 (1.46)		
17	**Single-lane two-way with tram (cluttered)**	Child	15.6	5.33 (3.34)	$\mathbf{t(72) = 1.49, p = 0.146}$	0.35
		Adult	20.5	4.09 (3.74)		

Note. Sample size varies per analysis due to the exclusion of participants who demonstrated erroneous responding. Clips with a non-significant difference between children and adults are presented in bold.
[a] Values represent seconds from the beginning of the window.
[b] Corrections for violations of homogeneity of variance have been made where appropriate. s = seconds.

Table 7
Final VR-PSCT Clips.

Hazard Perception Task		Gap Acceptance Task	
Clip Number	Roadway Environment	Clip Number	Roadway Environment
1	Single-lane two-way (cluttered)	1	Single-lane two-way (cluttered)
2	Single-lane two-way (uncluttered)	2	Single-lane two-way (uncluttered)
3	Single-lane one-way cycle heavy (uncluttered)	3	Single-lane one-way curved (cluttered)
4	Four-way intersection (cluttered)	4	Single-lane one-way cycle heavy (uncluttered)
5	Four-way intersection (uncluttered)	5	Four-way intersection (cluttered)
6	Multi-lane two-way (cluttered)	6	Four-way intersection (uncluttered)
7	Single-lane one-way (cluttered)	8	Multi-lane two-way (cluttered)
8	Single-lane one-way (uncluttered)	9	Single-lane one-way (uncluttered)
9	Pedestrian crossing (uncluttered)	10	Pedestrian crossing (cluttered)
10	Roundabout (uncluttered)	15	Single-lane two-way with refuge (cluttered)
13	Single-lane two-way with bike lane (uncluttered)	16	Single-lane two-way with refuge (uncluttered)
14	Single-lane two-way with refuge (uncluttered)		
16	Single-lane two-way with tram (uncluttered)		

Note. All clips included in the final version of the VR-PSCT can be downloaded from the Open Science Framework (https://osf.io/mk5xw/; McGuckian, 2024).

Table 8
Descriptive Statistics and t-tests for VR-PSCT Overall Scores.

	Child M (SD)	Adult M (SD)	*t*-test	Cohen's d
HP Response Times (s)	8.23 (2.88)	5.57 (2.12)	$t(49.7) = 4.32, p < 0.001$	1.08
GA Response Times (s)	4.56 (1.10)	2.75 (0.76)	$t(51.8) = 8.02, p < 0.001$	1.97
HP Accuracy	Median = 12 (IQR = 2.25)	Median = 13 (IQR = 1)	$U = 865$, z = 2.43, $p = 0.015$	–

principles used in the driving HP literature to develop a VR-PSCT which measured both HP and GA. We used principles of HP test development (Wetton et al., 2011) and standardised procedures of HP test development (Moran et al., 2019) to develop the VR-PSCT.

The application of these steps resulted in a final VR-PSCT with 13 clips (plus two control and two practice clips) for the HP task, and 11 clips (plus two control and two practice clips) for the GA task. All clips included in the final version of the VR-PSCT can be downloaded from the Open Science Framework (https://osf.io/mk5xw/; McGuckian, 2024). In line with principle one (Wetton et al., 2011), it was important to ensure that HP was measured as a separate construct to other factors such as GA and risk-taking propensity. To date research has focused on investigating either GA or HP in pedestrians, not both. GA has been the focus of a significant amount of research with this behaviour operationalised in several ways (Theofilatos et al., 2021), however there has been a comparative dearth of research on HP (Prabhakharan et al., 2024). In keeping with principle one, the task developed in this study effectively disentangles the skills of HP and GA. This is supported by good internal consistency demonstrated for the two constructs and evidence of divergent validity (i.e., weak non-significant correlations between the overall HP and GA response times). Being able to distinguish between these two skills is important as it allows for a thorough examination of these two skills, it will enable future research to better understand the predictors of these two independent behaviours, and it will support effective training of each independent road safety skill.

The research team made methodological choices to support the ecological validity of the VR-PSCT. First, the VR-PSCT comprises of a variety of scenario types which have previously been identified as problematic for pedestrians, including differing levels of visual clutter (Peel et al., 2023), different number of lanes (Kadali & Vedagiri, 2013), and different vehicle types (Rosenbloom et al., 2015). Second, 360-degree video footage was captured in line with principle three from Wetton et al. (2011). This was done from the average height of both children and adults, which maintained representative visual perspectives of the roadway and captured differences in visual cues (e.g., line of sight due to parked cars) which are known to affect street-crossing performance in these two groups (Meir, Oron-Gilad & Parmet, 2015). It would be important for future research to determine the validity of this task through comparisons between real life street-crossing performance and performance using the VR in the lab.

As predicted, children showed poorer street-crossing skill than adults for each individual clip and the overall scores on the VR-PSCT (see also Meyer et al., 2014; Rosenbloom et al., 2015). The majority of the clips included from the outset of this study demonstrated this pattern of performance, supporting the notion that comparing between adults and children aged seven to 11 years could be a viable option when selecting clips for inclusion in a pedestrian street-crossing task. That is, children represent a relatively inexperienced pedestrian population, while adults represent an experienced population. Therefore, the comparison between these two groups can be used in the same way that driving HP literature compares young inexperienced adult drivers to experienced adult drivers.

4.2. Limitations and future directions

In addition to the methodological rigour and principled approach used in this study, the VR-PSCT developed in this study has several strengths which improves on limitations of prior research. The combination of the use of a VR headset and 360-degree video footage of roadways allowed for participants to experience an immersive virtual environment. This virtual environment allows participants to freely explore the visual scene, as they would in real life, while maintaining the safety of participants and consistency of presentation of road crossing tasks for all participants. The use of this VR technology also enables other relevant information to be collected such as eye and head movements used to visually explore the roadway. Future research should make use of these forms of data to further validate the HP and GA tasks, and to gain an understanding of the importance of these visual exploratory behaviours for pedestrians.

A limitation to the validity of the VR-PSCT is that participants completed the task whilst seated and responded with a button press on their controller. Whilst these choices were made to ensure the safety of participants while wearing a VR headset, these are not natural pedestrian behaviours and therefore limits the ecological validity of the task. In particular this design means that walking speed of each individual participant was not captured which limits our ability to determine whether an individual would have safely made it across the street within the gap they had chosen. For GA tasks, future research should explore using more natural pedestrian behaviours, such as taking a step forward to indicate a road crossing decision, and walking on the spot to capture walking speed. This approach would require a consideration of 1) potential simulator sickness resulting from a mismatch between the visual environment and physical movement if 360-degree video footage is used, and 2) a method to accurately capture response data from participants. This latter point could be achieved by integrating additional technology such as a force plate or inertial measurement units, which can provide gait initiation data to indicate crossing decisions and measure walking speed.

This study focused on capturing only the behaviours of HP and GA. These were selected as pedestrians would need to identify potential hazards and select safe gaps to cross the road safely. This is an assumption, however, and there are other factors that may be considered by pedestrians when deciding to cross the road. We acknowledge that when making crossing decisions, pedestrians examine a variety of factors and may choose to engage in a number of different behaviours, and that this task does not capture all this complexity. We however have proposed a systematic approach to developing street-crossing tasks which could be applied to the development of tasks for a variety of different behaviours. This study aimed to capture roadway environments that prior research has demonstrated to impact on pedestrian safety. However, we were unable to include pedestrian crossing and signalised crossing situations in the VR-PSCT, despite these situations presenting risk for pedestrian safety (Di Stasi et al., 2014). Given that both driver and pedestrian behaviour is pre-determined by road rules in these situations (e.g., pedestrians have right of way at a pedestrian crossing, pedestrians must wait for signalised crossings) participants do not need to make a decision in the task and therefore these scenarios would not distinguish between good and poor GA ability. Engaging in jay walking behaviours at these locations increases the risk of pedestrians being involved in a crash (Besharati and Tavakoli Kashani, 2018), and therefore including these types of situations is valuable, however capturing jay walking in a testing environment is difficult as participants are more likely to engage in socially desirable behaviour when they know they are being tested. Methods to capture these important behaviours in street-crossing tasks warrants further investigation.

It is also important to consider the potential impact of the difference between children and adults in understanding the task requirements. We only identified child participants as erroneous responders, which could be a function of a lack of understanding of the task. We excluded those participants from data analysis so to not impact on the results, however this did reduce the sample size for children. We used a standardised script for both children and adults which provided operational definitions of the task in language that was suitable for children. Both children and adults were provided with an opportunity to practice both of the tasks, receive feedback and ask questions. These methods were employed to maximise understanding of the task requirements by children, however future research may want to employ more explicit checks of the task requirements with children to minimise the need for excluding cases.

Finally, this test was developed for an Australian context. Australia has different road designs to other countries and therefore it would be important for future research to investigate the roadway environments that would be most suitable for pedestrians in their context. The process for validation that was applied for the development of the VR-PSCT should be applied for the development of future street-crossing tasks in other contexts (e.g., other countries, rural areas, and roadway situations). It is important that future research attempts to replicate this validation process to confirm its applicability to street-crossing tasks for pedestrians. This method for validation could extend to tests which measure other pedestrian behaviours such time to collisions, safety decision, visual exploration, etc. To ensure confidence in the findings of pedestrian studies, it is important that future research provides evidence of validation of their street-crossing tasks.

4.3. Implications for road safety

Pedestrians account for approximately 23 % of all road fatalities (WHO, 2020). Road users, including pedestrians, are one of the core components of the safe systems framework (Elvik, 2023). If we are to achieve Vision Zero targets it is important that strategies for improvement that target road user behaviour have a strong evidential base. Applying best practice principles to the development of pedestrians tasks such as the VR-PSCT is important to provide confidence in our understanding of pedestrian behaviours. If we have confidence in our understanding of key pedestrian behaviours, and the potential factors that influence those behaviours, we are better positioned to develop targeted intervention strategies. Strategies could include training programs, or improvements in road infrastructure that can overcome some of the behavioural limitations identified through research using well-developed tasks.

To date there has been limited research in HP in pedestrians (Moran et al., 2019), and the work has often used fixed lab-based simulators. This study has demonstrated that VR can be utilised for the presentation of these tasks which means that we have high fidelity, portable, safe, relatively inexpensive, and user-friendly tools which can increase the potential for more researchers to engage in conducting pedestrian research. We need more researchers engaging in this work if we are to meet our targets. Furthermore, the portable nature of VR means that it is possible for tools such as the VR-PSCT to include participants who might otherwise be unable to attend a lab session at a university. This could include participants such as individuals with disabilities and older adults with mobility issues. Increasing the representation of our samples through using portable research tools will increase the generalisability of findings, and act to ensure that any interventions developed are fit for all persons.

4.4. Conclusion

It is important that road safety research maintains rigorous scientific processes when investigating the behaviour of road users. By adopting principles and established approaches from driving HP literature, we were able to successfully apply these principles to the development of the VR-PSCT, a pedestrian road safety task that uses 360-video to present representative road cross situations using a VR headset. The VR-PSCT discriminates between children and adults based on their performance in both HP and GA tasks, which have been shown to be distinct pedestrian skills. Our detailed method may provide a model for pedestrian task development in future road safety research. Applying best practice principles to the development of tasks such as the VR-PSCT, means that these tasks can be used to investigate important predictors of pedestrian road crossing behaviour, and inform training interventions and education programs.

CRediT authorship contribution statement

Joanne M. Bennett: Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Formal analysis, Conceptualization. **Thomas B. McGuckian:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis, Conceptualization. **Nathan Healy:** Writing – review & editing, Investigation, Data curation. **Nikki Lam:** Writing – review & editing, Investigation, Data curation. **Ralph Lucas:** Writing – review & editing, Investigation, Data curation. **Kathleen Palmer:** Writing – review & editing, Investigation, Data curation. **Robert G. Crowther:** Writing – review & editing, Software. **David A. Greene:** Writing – review & editing, Supervision. **Peter Wilson:** Writing – review & editing, Supervision. **Jonathan Duckworth:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Asher, L., Aresu, M., Falaschetti, E., Mindell, J., 2012. Most older pedestrians are unable to cross the road in time: a cross-sectional study. Age Ageing 41 (5), 690–694. https://doi.org/10.1093/ageing/afs076.

Besharati, M.M., Tavakoli Kashani, A., 2018. Which set of factors contribute to increase the likelihood of pedestrian fatality in road crashes? Int. J. Inj. Contr. Saf. Promot. 25 (3), 247–256. https://doi.org/10.1080/17457300.2017.1363781.

Crundall, D., Van Loon, E., Baguley, T., Kroll, V., 2021. A novel driving assessment combining hazard perception, hazard prediction and theory questions. Accid. Anal. Prev. 149, 105847. https://doi.org/10.1016/j.aap.2020.105847.

Deb, S., Carruth, D.W., Sween, R., Strawderman, L., Garrison, T.M., 2017. Efficacy of virtual reality in pedestrian safety research. Appl. Ergon. 65, 449–460. https://doi.org/10.1016/j.apergo.2017.03.007.

Di Stasi, L.L., Megías, A., Cándido, A., Maldonado, A., Catena, A., 2014. The influence of traffic signal solutions on self-reported road-crossing behavior. Span. J. Psychol. 17, E103.

Dommes, A., Cavallo, V., 2011. The role of perceptual, cognitive, and motor abilities in street-crossing decisions of young and older pedestrians. Ophthalmic Physiol. Opt. 31 (3), 292–301. https://doi.org/10.1111/j.1475-1313.2011.00835.x.

Elvik, R., 2023. What would a road safety policy fully consistent with safe system principles mean for road safety? Accid. Anal. Prev. 193, 107336. https://doi.org/10.1016/j.aap.2023.107336.

Haleem, K., Alluri, P., Gan, A., 2015. Analyzing pedestrian crash injury severity at signalized and non-signalized locations. Accid. Anal. Prev. 81, 14–23. https://doi.org/10.1016/j.aap.2015.04.025.

Hill, A., Horswill, M.S., Whiting, J., Watson, M.O., 2019. Computer-based hazard perception test scores are associated with the frequency of heavy braking in everyday driving. Accid. Anal. Prev. 122, 207–214. https://doi.org/10.1016/j.aap.2018.08.030.

Horswill, M.S., 2016. Hazard perception in driving. Curr. Dir. Psychol. Sci. 25 (6), 425–430. https://doi.org/10.1177/0963721416663186.

Horswill, M.S., McKenna, F.P., 2004. Drivers' hazard perception ability: situation awareness on the road. A cognitive approach to situation awareness: Theory and Application 155–175.

Horswill, M.S., Hill, A., Silapurem, L., Watson, M.O., 2021. A thousand years of crash experience in three hours: an online hazard perception training course for drivers. Accid. Anal. Prev. 152, 105969. https://doi.org/10.1016/j.aap.2020.105969.

Ishaque, M.M., Noland, R.B., 2008. Behavioural issues in pedestrian speed choice and street crossing behaviour: a review. Transp. Rev. 28 (1), 61–85. https://doi.org/10.1080/01441640701365239.

Jackson, L., Chapman, P., Crundall, D., 2009. What happens next? Predicting other road users' behaviour as a function of driving experience and processing time. Ergonomics 52 (2), 154–164. https://doi.org/10.1080/00140130802030714.

Kadali, B.R., Vedagiri, P., 2013. Effect of vehicular lanes on pedestrian gap acceptance behaviour. Procedia Soc. Behav. Sci. 104, 678–687. https://doi.org/10.1016/j.sbspro.2013.11.162.

Kadali, B.R., Vedagiri, P., 2020. Role of number of traffic lanes on pedestrian gap acceptance and risk taking behaviour at uncontrolled crosswalk locations. J. Transp. Health 19, 100950. https://doi.org/10.1016/j.jth.2020.100950.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G., 1993. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol. 3 (3), 203–220. https://doi.org/10.1207/s15327108ijap0303_3.

Lassarre, S., Papadimitriou, E., Yannis, G., Golias, J., 2007. Measuring accident risk exposure for pedestrians in different micro-environments. Accid. Anal. Prev. 39 (6), 1226–1238. https://doi.org/10.1016/j.aap.2007.03.009.

Leung, K.Y., Loo, B.P., Tsui, K.L., So, F.L., Fok, E., 2021. To cross or not to cross: a closer look at children's decision-making on the road. Transp. Res. A Policy Pract. 149, 1–11. https://doi.org/10.1016/j.tra.2021.04.011.

Levulytė, L., Baranyai, D., Sokolovskij, E., Török, Á., 2017. Pedestrians' role in road accidents. Int. J. Traffic Transp. Eng. 7 (3), 328–341.

Li, H., Wu, D., Graham, D.J., Sze, N.N., 2020. Comparison of exposure in pedestrian crash analyses: a study based on zonal origin-destination survey data. Saf. Sci. 131, 104926. https://doi.org/10.1016/j.ssci.2020.104926.

Liu, Y.C., Tung, Y.C., 2014. Risk analysis of pedestrians' road-crossing decisions: effects of age, time gap, time of day, and vehicle speed. Saf. Sci. 63, 77–82. https://doi.org/10.1016/j.ssci.2013.11.002.

McGuckian, T. (2024). Development of a Virtual Reality Pedestrian Street Crossing Task - Video Assets. Retrieved from osf.io/mk5xw.

Meir, A., Oron-Gilad, T., Parmet, Y., 2015. Are child-pedestrians able to identify hazardous traffic situations? Measuring their abilities in a virtual reality environment. Saf. Sci. 80, 33–40. https://doi.org/10.1016/j.ssci.2015.07.007.

Meyer, S., Sagberg, F., Torquato, R., 2014. Traffic hazard perception among children. Transport. Res. F: Traffic Psychol. Behav. 26, 190–198. https://doi.org/10.1016/j.trf.2014.07.007.

Montufar, J., Arango, J., Porter, M., Nakagawa, S., 2007. Pedestrians' normal walking speed and speed when crossing a street. Transp. Res. Rec. 2002 (1), 90–97. https://doi.org/10.3141/2002-12.

Moran, C., Bennett, J.M., Prabhakharan, P., 2019. Road user hazard perception tests: a systematic review of current methodologies. Accid. Anal. Prev. 129, 309–333. https://doi.org/10.1016/j.aap.2019.05.021.

Moran, C., Bennett, J.M., Prabhakharan, P., 2020. The relationship between cognitive function and hazard perception in younger drivers. Transport. Res. F: Traffic Psychol. Behav. 74, 104–119. https://doi.org/10.1016/j.trf.2020.08.008.

New South Wales Government. (2023). Design of roads and streets: A guide to improve the quality of roads and streets in NSW. https://www.movementandplace.nsw.gov.au/sites/default/files/guides/pdf/design-of-roads-and-streets.pdf.

Office of Road Safety. (2021). *National Road Safety Strategy 2021-2023 Fact Sheet.* https://www.roadsafety.gov.au/nrss/fact-sheets/vulnerable-road-users.

Oxley, J.A., Ihsen, E., Fildes, B.N., Charlton, J.L., Day, R.H., 2005. Crossing roads safely: an experimental study of age differences in gap selection by pedestrians. Accid. Anal. Prev. 37 (5), 962–971. https://doi.org/10.1016/j.aap.2005.04.017.

Peel, M. A., Tigue, M. B., Amores, K. A., & Palmer, E. M. (2023, October). Effects of Traffic Clutter and Road Type on Road Hazard Detection. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* https://doi.org/10.1177/21695067231192546.

Prabhakharan, P., Bennett, J.M., Hurden, A., Crundall, D., 2024. The efficacy of hazard perception training and education: a systematic review and meta-analysis. Accid. Anal. Prev. 202, 107554. https://doi.org/10.1016/j.aap.2024.107554.

Rosenbloom, T., Mandel, R., Rosner, Y., Eldror, E., 2015. Hazard perception test for pedestrians. Accid. Anal. Prev. 79, 160–169. https://doi.org/10.1016/j.aap.2015.03.019.

Rothman, L., Howard, A.W., Camden, A., Macarthur, C., 2012. Pedestrian crossing location influences injury severity in urban areas. Inj. Prev. 18 (6), 365–370. https://doi.org/10.1136/injuryprev-2011-040246.

Scialfa, C.T., Borkenhagen, D., Lyon, J., Deschênes, M., Horswill, M., Wetton, M., 2012. The effects of driving experience on responses to a static hazard perception test. Accid. Anal. Prev. 45, 547–553. https://doi.org/10.1016/j.aap.2011.09.005.

Shen, Z., Ma, J., Wang, N., 2023. Development and validation of a video-based assessment tool for children's street-crossing safety. J. Transp. Health 33, 101716. https://doi.org/10.1016/j.jth.2023.101716.

SWOV. (2020). *Pedestrians.* SWOV fact sheet, July 2020. SWOV, The Hague.

Tapiro, H., Oron-Gilad, T., Parmet, Y., 2020. Pedestrian distraction: the effects of road environment complexity and age on pedestrian's visual attention and crossing behavior. J. Saf. Res. 72, 101–109. https://doi.org/10.1016/j.jsr.2019.12.003.

Theofilatos, A., Ziakopoulos, A., Oviedo-Trespalacios, O., Timmis, A., 2021. To cross or not to cross? Review and meta-analysis of pedestrian gap acceptance decisions at midblock street crossings. J. Transp. Health 22, 101108. https://doi.org/10.1016/j.jth.2021.101108.

Unterberger, P., & Johnston., P. (2015). *Pedestrians and Road Safety*. Department of Infrastructure and Regional Development. https://www.bitre.gov.au/sites/default/files/is_070.pdf.

Wetton, M.A., Hill, A., Horswill, M.S., 2011. The development and validation of a hazard perception test for use in driver licensing. Accid. Anal. Prev. 43 (5), 1759–1770. https://doi.org/10.1016/j.aap.2011.04.007.

World Health Organization. (2013). *Pedestrian safety: A road safety manual for decision-makers and practitioners.* https://www.who.int/publications/i/item/pedestrian-safety-a-road-safety-manual-for-decision-makers-and-practitioners.

World Health Organization. (2020). *Road Traffic Injuries.* https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

Yıldırım, Ö.C., Celik, E., 2023. Understanding pedestrian behavior and spatial relations: a pedestrianized area in Besiktas Istanbul. Front. Architectural Res. 12 (1), 67–84. https://doi.org/10.1016/j.foar.2022.06.009.