Transportation Research Part F: Psychology and Behaviour 113 (2025) 586–608

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

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

Head movements predict pedestrian crossing decisions earlier than walking speed☆

Max Theisen [a,*], Melina Bergen [a], Wolfgang Einhäuser [b], Caroline Schießl [a]

[a] Institute of Transportation Systems, German Aerospace Center (DLR), Lilienthalpl. 7, 38108 Braunschweig, Germany
[b] Institute of Physics, Chemnitz University of Technology, Reichenhainer Str. 70, 09126 Chemnitz, Germany

ARTICLE INFO

Keywords:
Pedestrian-vehicle interaction
Pedestrian crossing
Head movements
Prediction
Virtual reality
Omnidirectional treadmill

ABSTRACT

Understanding the behaviour of pedestrians when crossing roads is crucial to improving safety, as pedestrians are the most vulnerable road users, but also the most difficult to predict. Previous research has examined how pedestrians, who intend to cross a road, assess vehicle speed and distance for gap acceptance. However, little attention has been given to the role of head movements as predictor for crossing decisions. Besides by the crossing intention, head movements are also influenced by the approach process to the kerb and the potential traffic directions, but both factors are rarely studied. Here, we compare the predictive power of head movements and walking speed for crossing decisions by examining how pedestrians approach a crossing location while simultaneously looking for crossing opportunities. We used an omnidirectional treadmill combined with a head-mounted display to investigate the crossing behaviour of 36 participants. Our results demonstrate that in the first seconds after pedestrians notice the car, the head-turning frequency is a stronger predictor for the crossing decision than the walking speed. As time passes, the walking speed becomes more useful than the head information and reaches a perfect predictive power shortly before pedestrians cross the kerb. These findings suggest that head orientation could be used as a predictive feature in the initial phase of the crossing and walking speed in the later phase. Ultimately, this research highlights the need for future studies that investigate pedestrian behaviour under more realistic conditions, providing insights relevant to behaviour prediction models for pedestrian protection in automated driving.

1. Introduction

Pedestrians have the highest injury risk in a collision with a vehicle and are therefore considered as the most vulnerable among all road users (Lubbe et al., 2022). Despite a general increase in road safety over the last decades, in 2021 approximately 250,000 pedestrians were killed in road traffic related accidents worldwide (WHO, 2023). Understanding pedestrian behaviour in road crossing scenarios is especially critical, as this is where they most often come into conflict with vehicles. Analysing and predicting how pedestrians act when crossing roads can yield valuable insights for traffic safety measures (e.g., the design of pedestrian crossings), policy

☆ This article is part of a special issue entitled: 'Age of vehicle automation' published in Transportation Research Part F: Psychology and Behaviour.

* Corresponding author.
E-mail addresses: max.theisen@dlr.de (M. Theisen), melina.bergen@dlr.de (M. Bergen), wolfgang.einhaeuser-treyer@physik.tu-chemnitz.de (W. Einhäuser), caroline.schiessl@dlr.de (C. Schießl).

https://doi.org/10.1016/j.trf.2025.05.015
Received 28 December 2024; Received in revised form 7 May 2025; Accepted 13 May 2025
Available online 31 May 2025

formulation and technological solutions such as autonomous vehicles (AVs). Studies have attributed a significant proportion of traffic accidents to human error – over 90%, according to (Treat et al., 1979) and (Singh, 2015) – though these estimates include cases where human error was one of multiple factors or only a contributing factor. As a result, AVs have been seen as a potential means of improving road safety (Litman, 2020), particularly for vulnerable road users like pedestrians (Fagnant and Kockelman, 2015), though this remains an ongoing challenge. As AVs become more common, they must be programmed to predict and adapt to the behaviour of pedestrians. One important use-case is to predict whether a pedestrian is about to cross the road in front of an AV, a topic that has been studied intensively over the past years (Keller and Gavrila, 2013; Koehler et al., 2013; Kooij et al., 2014; Rasouli et al., 2017b). The prediction of a pedestrian's crossing intent would in consequence force the AV to change its own actions, for example, by yielding to the pedestrian to maximise safety. The earlier the AV is able to predict the pedestrian's intentions, the earlier it can plan its own trajectory accordingly and therefore create a more time-, energy- and safety-efficient interaction.

1.1. Predicting pedestrian intent

Forecasting pedestrian behaviour requires detection, tracking, and prediction algorithms. Pedestrian detection is typically achieved using manually defined features (e.g., HOG/SVM (Dalal and Triggs, 2005)) or deep learning-based methods (e.g., YOLO (Redmon et al., 2016)), with the latter offering greater speed and accuracy (Brunetti et al., 2018; Cao et al., 2021). In the context of road crossings, various models have been used to classify pedestrian actions (crossing vs. stopping initiation), including Kalman filters and Gaussian process dynamical models (Keller and Gavrila, 2013; Koehler et al., 2013; Mínguez et al., 2018). While technical systems have improved in accuracy over time, human observers still outperform them in predicting pedestrian intent well before the event (Keller and Gavrila, 2013). For comprehensive reviews, see (Ridel et al., 2018) for pedestrian intention estimation and (Korbmacher and Tordeux, 2022) for trajectory prediction.

1.2. Head movements as predictors of pedestrian crossing decisions

Notably, information on head orientation was integrated into pedestrian prediction models (Kooij et al., 2014; Schulz and Stiefelhagen, 2015). (Rehder et al., 2014) were able to estimate pedestrians' head orientation using a HOG/SVM combination with a mean absolute error of 19°, proving a real-time recognition of head orientation from camera images. (Flohr et al., 2015) achieved similar results of 21°/19° using a probabilistic framework for the joint estimation of pedestrian head and body orientation. (Kloeden et al., 2014) equipped the heads of different pedestrians, walking on the sidewalk parallel to a road, with inertial measurement units. Analysing the change of head orientation, the authors were able to detect if someone would turn to cross the road (yielding more head turns) on average 3.8 s before the pedestrian would deviate more than 20° from walking straight. Additionally, they analysed whether the stopping of pedestrians could be predicted when they walked perpendicular to the road, approaching it head-on. Here the authors concluded that in the case of stopping (and not crossing) the pedestrians no longer look to both sides, but only fixate the approaching vehicle. In a test track study in Japan, (Hamaoka et al., 2013) investigated head-turning of pedestrians crossing a crosswalk from both directions in the presence of interfering left- and right-turning cars. Head-turning frequency increased towards the entry of the crosswalk from both sides, most prominently when the conflict point was in the near lane. If the conflict point was in the far lane from the pedestrians' perspective, head-turning frequency peaked again when reaching this point. Similarly, (Hassan et al., 2005) reported an increased head-turning behaviour in the last second before crossing the street for different types of intersections and for people with and without visual impairment.

1.3. Real-world observations of pedestrian head movements

Additionally, in naturalistic observations head orientation is often used as a proxy for gaze behaviour and visual attention. (Pipkorn et al., 2024) found that bidirectional driver-pedestrian gaze plays a role in pedestrian decision-making, suggesting that mutual gaze can be a factor in crossing interactions. Similarly, (Rasouli et al., 2017a) emphasised the role of head movements in signalling crossing intent. However, while gaze may influence interactions in some cases, research suggests that pedestrians often do not depend on eye contact to decide when to cross. In fact, pedestrians rarely use explicit cues from drivers (e.g., honking, flashing headlights, or gestures), and instead rely on implicit kinematic cues of the vehicle (Lee et al., 2021). The observation that most pedestrians begin crossing before they can even see the driver's face (AlAdawy et al., 2019) further challenges the assumption that they depend on eye contact to the driver. These findings reinforce the notion that pedestrians primarily make crossing decisions based on vehicle kinematics rather than on mutual gaze interaction with the driver. While pedestrians may not rely on gaze, findings show that drivers are indeed influenced by pedestrian gaze behaviour: (Guéguen et al., 2015) found that eye contact significantly increased the likelihood of drivers stopping, and (Ren et al., 2016) demonstrated that eye contact alters drivers' deceleration patterns. So, in addition to perceiving information from the dynamic vehicle cues for the crossing decision, head movements toward the vehicle could also serve to signal the pedestrian's own intention to cross to the driver.

1.4. Controlled laboratory studies of pedestrian head movements

Many different cues have been believed to signal the pedestrian's crossing intention. Using videos from natural traffic scenes (Schmidt et al., 2008) asked participants which cues they used to judge the pedestrian's future crossing decision. Results showed only 16% mentioned traffic (e.g., velocity of vehicles), while 75% of statements were related to the pedestrian. More specifically, 28% of

total comments were related to head movements, 24% to locomotion and and 10% to leg and feet movement. In a follow-up study (Schmidt and Faerber, 2009) investigated whether masking different aspects of the traffic scene (e.g., pedestrian's head, legs, complete body [including the head] or surrounding traffic scene) influenced participants' ability to predict pedestrians' crossing decisions. Only in the condition where the pedestrians' entire body was masked, did the observers make significantly more mistakes, but not when only the head or legs were masked. This implies that only the trajectory might not be enough to predict the pedestrian's crossing intention accurately, rather the movement of the lower (legs plus torso) or upper body (head plus torso) provide important additional information to assess the crossing decision. In a laboratory motion tracking study (Kalantarov et al., 2018) were able to quantify the crossing initiation process (from stand-still) for different body parts. The authors showed that crossing started with the forward motion of head and shoulder (−0.8 s before ankle), followed by hip, wrist and elbow (−0.6 s), knee (−0.3 s) until finally the ankle is lifted forward to make the first step onto the road.

Clues that head movements may provide valuable information about the pedestrian's crossing intention have long been known from traffic observations (see Subsection 1.2) but few studies have tested this in controlled laboratory settings, (Yang et al., 2024; Lyu et al., 2024) being recent exceptions. (Yang et al., 2024) measured pedestrians' head movements inside a cave automatic virtual environment (CAVE) for an AV (that yielded in 75% of cases) approaching from the right side of the pedestrian, who was always stationary at the start of the trial. They found an immediate increase in absolute head-turning frequency from the start of the trial until 7 s before crossing initiation (peaking at around 60°/s), which is likely due to a pre-orientation of pedestrians' heads to the right side (the only direction from which vehicles could approach). Next, from 6 to 1 s before crossing initiation, the absolute head-turning frequency was approximately constant around 20°/s, showing that pedestrians' were likely monitoring the approaching vehicle and were accumulating evidence for the crossing-decision from changing parameters like time-to-arrival, distance, acceleration and eHMI signals (Pekkanen et al., 2022). Finally, from 1 s before to 1.5 s after crossing initiation the absolute head-turning frequency was increasing again to around 70°/s, and was staying on a higher level of around 50°/s also during crossing. In another study from the same CAVE laboratory, (Lyu et al., 2024) investigated whether pedestrians' head movements differ between crossing and non-crossing manoeuvres (starting from stand-still). They found that from 2 s prior to crossing initiation the absolute head-turning frequency increased more rapidly and was significantly higher in trials that resulted in a crossing compared to non-crossing manoeuvres and hence could serve as an early indicator of road crossing intentions.

1.5. Contribution of this study

In the present study we investigate the predictability of pedestrians' road crossing behaviour. We extend previous studies in the following respect: First, previous studies typically assume that pedestrians are stationary at the beginning of the crossing decision process (Kalantarov et al., 2018; Sobhani and Farooq, 2018; Kooijman et al., 2019; Schneider et al., 2022; Yang et al., 2024; Lyu et al., 2024). However, observations from real-traffic (Gorrini et al., 2018) suggest that the crossing decision-making is already happening while approaching the kerb, and pedestrians who decide to cross the road do not always (unnecessarily) stop at the kerb before crossing (Rasouli et al., 2017b). Rather, in most cases, the participant's decision whether or not to cross the road manifests itself in either accelerating or decelerating before reaching the kerb. According to (Gorrini et al., 2018), the time course of pedestrians' walking speed during the crossing process can be divided into three phases:

I. Approaching: walking speed is approximately constant
II. Appraising: walking speed decreases as pedestrians decelerate to evaluate dynamic cues of approaching vehicles (e.g., speed, distance)
III. Crossing/Stopping: walking speed either increases as pedestrians accelerate to cross the road or walking speed further decreases as pedestrian decide not to cross

Therefore results from laboratory studies using a standing start might not generalise entirely to real-world behaviour. We address this by studying how pedestrians approach a crossing location while looking for crossing opportunities at the same time. Technically this is made possible by using a treadmill system, the "Omnideck" (Brinks and Bruins, 2016), an omnidirectional treadmill that enables continuous walking in Virtual Reality (VR).

Secondly, many controlled laboratory studies only focus on traffic approaching from one side (Kooijman et al., 2019; Yang et al., 2024; Lyu et al., 2024; Sobhani and Farooq, 2018; Schneider et al., 2022) or using single-lane roads (Kalantarov et al., 2018; Yang et al., 2024; Lyu et al., 2024; Sobhani and Farooq, 2018; Schneider et al., 2022) limiting the natural demand for visual search and safety assurance through head-turns. In the real world, pedestrians walk towards the kerb and have traffic from both directions, but both are rarely investigated in experimentally well-controlled studies. At the same time, real-world studies (unavoidably) leave many parameters uncontrolled (traffic density, vehicle appearance, vehicle speed, vehicle TTA, etc.). Here we use virtual reality to combine both, high ecological validity and high experimental control. With our experimental design we provide repeated measures for the same pedestrian as well as experimental control over the traffic situation. Thereby we aim to measure observable behavioural cues that may signal a pedestrian's crossing decision. Specifically, we hypothesise that, in addition to walking speed/trajectory, head movements play a role in this process, as they help gather visual information relevant to the decision-making process.

2. Methods

2.1. Participants

Thirty-six participants (11 female, 25 male) aged between 19 and 46 years took part in the experiment (M = 28.7 years, SD = 7.5 years). Walking on the omnidirectional treadmill requires good balance and bears a risk of falling into the safety harness. In order to minimise the risk of injury the age limit was set rather conservatively to 50 years. Exclusion criteria were:

- age under 18 or over 50 years old
- body mass over 120 kg (device limitation)
- known propensity for simulator sickness (or any sign of occurrence during experiment which did not occur in this study)
- impaired vision that prevents use of the VR headset (eye tracking), blindness, conjunctivitis
- pregnancy (to minimise the risk of injury in the event of the belt safety system being activated)
- back problems or injury, for example, herniated disk
- migraine
- cardiovascular diseases
- neurological conditions, for example, photosensitive epilepsy
- impaired balance or oculomotor system

As detailed in the Results section (3.2), data of four of the 36 participants had to be excluded from the main analysis as their crossing behaviour was consistent in all trials and independent of condition or because of implausible walking speed data due to technical error. Study participants were recruited via the participant pool of the Institute of Transportation Systems of the DLR and can be assumed to be experienced with traffic-psychology-related experiments. All participants took part voluntarily. The study took approximately 90 min and the participants received 12 € per hour as a compensation for their participation. An ethical review application was submitted to the ethics commission of the DLR. The ethics commission stated that there is no relevance for an assessment and decided to waive the requirement of an in-depth ethical review for this study.

2.2. Apparatus

The experiment was conducted on an omnidirectional walking simulator, which allows the participants to move freely in any direction. Omnidirectional treadmills provide the most realistic walking experience of all current walking simulators (Novacek and Jirina , 2022). The Omnideck (W5 Omnifinity AB, Växjö/Sweden) (see Fig. 1) is a motorised treadmill with a scaleable diameter, which in case of the MoSAIC – Laboratory for multi-user simulation at the Institute of Transportation Systems, Braunschweig – is 3.7 m wide. It consists of 16 wedges of motorised spinning tubes, that counter the movements of the user and transport the user to the centre of the walking area (Novacek and Jirina , 2022; Brinks and Bruins, 2016). The speed of the rollers depends on the radial position of the user. Specifically, we used the "progressive" base speed algorithm in the Omnitrack API (the software responsible for controlling the Omnideck) so as the user gets further to the edges, the rollers spin exponentially faster, moving the user back to the centre. In the centre of the Omnideck there is a static platform with a diameter of 0.5 m, from which the participants start walking. The participants wore a safety harness which is attached to the ceiling in the middle of the treadmill. The virtual environment was presented using a Vive Pro

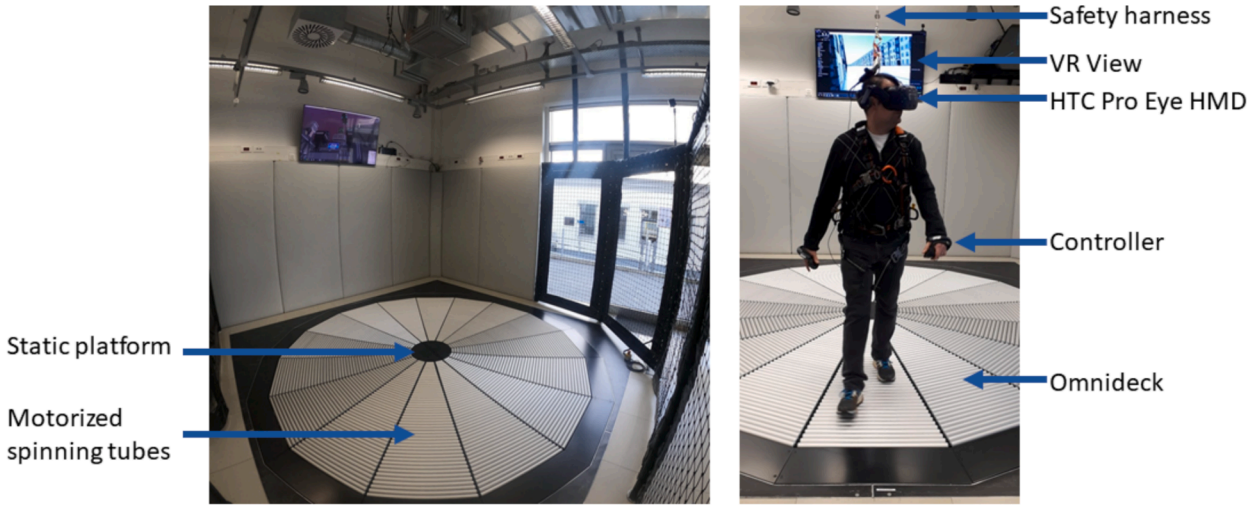


Fig. 1. Setup of Omnideck in the MoSAIC – Laboratory for multi-user simulation at the Institute of Transportation Systems, Braunschweig.

Eye (HTC Coperation, Taoyuan/Taiwan) with a resolution of 1440 x 1600 pixels per eye, a refresh rate of 90 Hz and a field-of-view specified as 110° by the manufacturer. The head-mounted display (HMD) had integrated eye tracking. In addition to the HMD, the participants received controllers which were strapped to their hands and tracked the movement of the hands. The sampling rate of all measured variables (e.g., walking speed, head rotation, etc.) was 50 Hz.

2.3. Materials

The virtual environment consisted of a virtual city with a long alley surrounded by houses, which was perpendicular to 10 two-lane roads (Fig. 2). The scenario was aligned such that the pedestrians walked in negative y-direction and the streets ran in x-direction. On each of the roads either one car approached from the left (45% of cases; Fig. 2a) or the right side (45% of cases; Fig. 2b) or there was no car at all (10% of cases). The cars started to drive when the pedestrian activated an invisible trigger point at the end of the alley (5.5 m away from the kerb; Fig. 2c) leading to that road, which was activated when the pedestrian walked through them. This ensured that the gap sizes were the same across participants, regardless of how fast they moved through the scenario. At the point of trigger activation the car was not yet visible at that moment. The car used in the scenario is an autonomously driving, white *BMW i3* (i.e., no driver was behind the wheel). The car was driving at a constant velocity of 30 km/h, 40 km/h or 50 km/h and did not react to the pedestrians' movements in any way. Regarding the starting point we used different time-to-arrival conditions: 11 s, 13 s or 15 s for cars approaching from the left and 14 s, 16 s or 18 s for cars approaching from the right, as it took approximately 3 s to cross the first lane. Each participant completed a total of 40 road crossing trials, divided into four blocks of 10 trials each. Nine of these were regular trials with an approaching vehicle, while one was a baseline trial without any vehicle. Each of the three velocity levels and six TTA levels was presented twice per participant. This resulted in a total of 36 trials per participant in which the participant interacted with a vehicle plus four baseline trials. The order of the 18 velocity-TTA-lane combinations and the positioning of the baseline trials were counterbalanced across participants using a Latin square design, ensuring that each block contained exactly one baseline trial. Collisions with the car did not occur, but if they had, they would have been non-interactive. We measured the controller-, head- and eye movements, the walking speed and trajectory of the participants. The virtual environment was implemented using the 3D computer graphics game engine, Unreal Engine 4.

2.4. Procedure

After consenting to participation, participants filled out a questionnaire about their demographic data and then were familiarised with the simulator and secured with the harness. Additionally, the eye-tracking was calibrated. The calibration was repeated before each block (or every time the participant had taken off the VR headset). The participants were instructed to walk along the alley in the virtual city. They were asked to behave as they would in real life when they reached a road, and safely cross the road without running

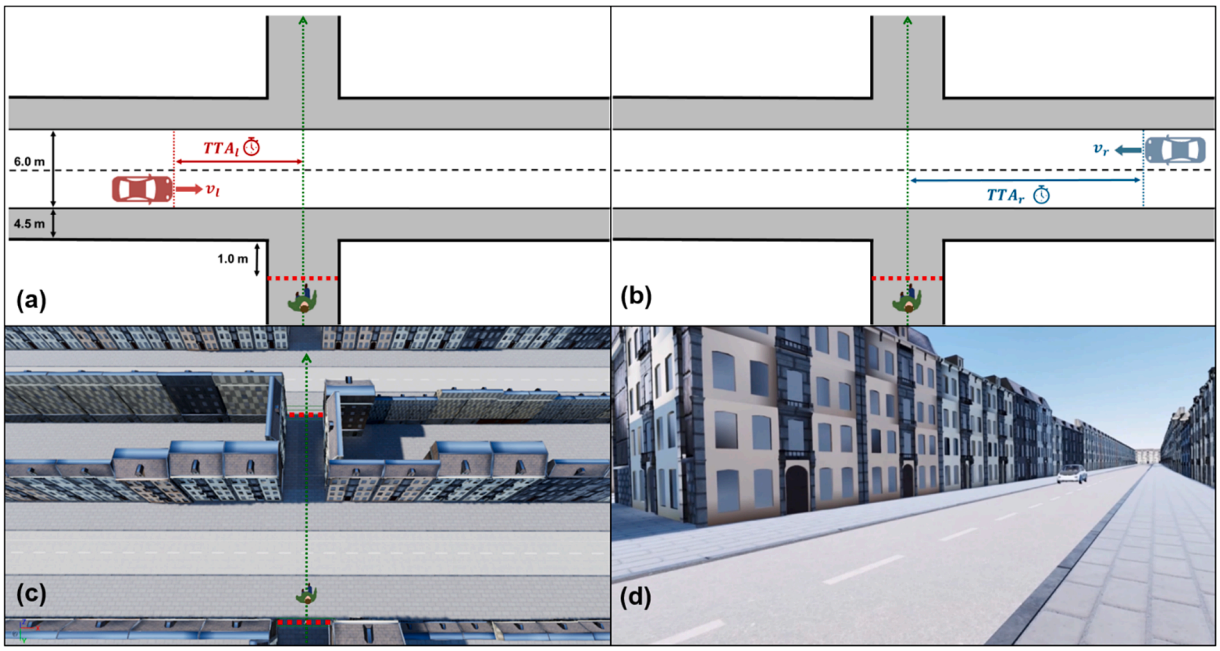


Fig. 2. Sketch of scenario with vehicle approaching from left (a) and right side (b), which both were equally frequent (45%). In 10% of trials no car was approaching at all (baseline). The virtual scenario is depicted from bird's view (c) and first person (d) perspectives. Dashed red lines in panels (a) through (c) depict the trigger point (invisible to the participant). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

or unnecessary delays. Before starting the actual experiment, the participants completed a training in order to get used to walking on the simulator and to make sure that the participants understood the task. The virtual environment for the training was the same as for the experimental trials. The only difference is that the first two roads were guaranteed to be free of approaching vehicles. When the participant felt sufficiently safe and was able to walk naturally on the simulator, the actual experiment began. The participants completed four blocks, consisting of 10 roads each. In between the blocks the participants got a 5-minute break. After the completion of the four blocks, the participants received a debriefing, in which the purpose of the study was explained in detail and participants were given the opportunity to express comments or questions.

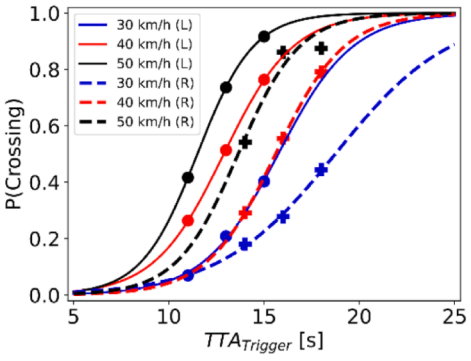
3. Results

3.1. Gap acceptance

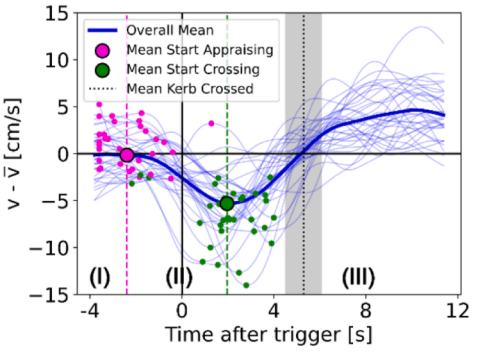
As a first analysis, we assess the behavioural data, which is typically also investigated in other laboratory studies to check the extent to which we replicate earlier studies in this respect. One important traffic safety measure is gap acceptance, referring to the temporal or spatial gaps to approaching vehicles that pedestrians assess to be (just) big enough to cross safely. Trivially, we find that pedestrians have a higher crossing probability for vehicles that were triggered with a bigger time gap (Cohen et al., 1955). We find that the crossing probability is reduced if the vehicle is approaching from the right (far) lane, compared to the approach from the left (near) lane (cf. Fig. 3(a)). This is because the effective time gap is smaller for the vehicles approaching from the right since the pedestrian needs additional time to cross the near lane. While the pattern of crossing probabilities matches the literature, the 50% threshold time gaps that we measure (11–13 s for cars driving 50–40 km/h) are much bigger compared to previous studies reporting 4–7 s (Lobjois and Cavallo, 2007; Petzoldt, 2014; Tian et al., 2022) likely because those studies refer to pedestrians standing at the kerb. In contrast, $TTA_{Trigger}$ in this study refers to the earliest possible time where pedestrians could see the vehicle, while they are still 5.5 m away from the kerb. If we subtract the on average 5–6 s time it took for pedestrians to reach the kerb (cf. Fig. 3(a)) the results become comparable to those earlier studies. We also replicate that pedestrians show a higher crossing probability for faster cars given a constant TTA, now in a walking treadmill simulator study combined with VR. This finding seems to be robust independent of experimental environment in online (Theisen et al., 2024), laboratory screen (Oxley et al., 2005; Petzoldt, 2014), laboratory CAVE (Lobjois and Cavallo, 2007) and real-world test-track studies (Schmidt and Faerber, 2009). To validate our results we used a logistic regression with a binomial distribution and a logit link function, which confirmed that higher TTA significantly increased the probability of crossing for both left-side-first-lane ($\beta = 0.58$, $p < 0.001$) and right-side-second-lane vehicles ($\beta = 0.46$, $p < 0.001$). Higher vehicle speed also increased crossing probability, with significant effects in both models (left-side-first-lane: $\beta = 0.12$, $p < 0.001$; right-side-second-lane: $\beta = 0.11$, $p < 0.001$). The pseudo R^2 values (Cox and Snell, 1989) were 26.8% (left-side-first-lane) and 22.5% (right-side-second-lane. An interaction between TTA and speed was not significant ($p > 0.05$).

3.2. Walking speed

To check whether we were successful in creating a natural approach behaviour of pedestrians to the kerb (instead of starting from a standing position), we investigate the pedestrians walking speed: In trials where pedestrians decided to cross in front of the oncoming



(a) Aggregated crossing probability of pedestrians for different car speed and TTA conditions supplemented by a sigmoidal trend as a guide-to-the-eye



(b) Average deviation from mean walking speed per participant when crossing (thin blue curves) and overall mean (thick blue curve). Black dotted line shows mean time when kerb is crossed, shaded area shows SD. Crossing phases according to [Gorrini et al., 2018]: (I) Approaching, (II) Appraising, (III) Crossing. See text for delineation of the phases.

Fig. 3. Crossing probability (left) and walking speed data when crossing (right).

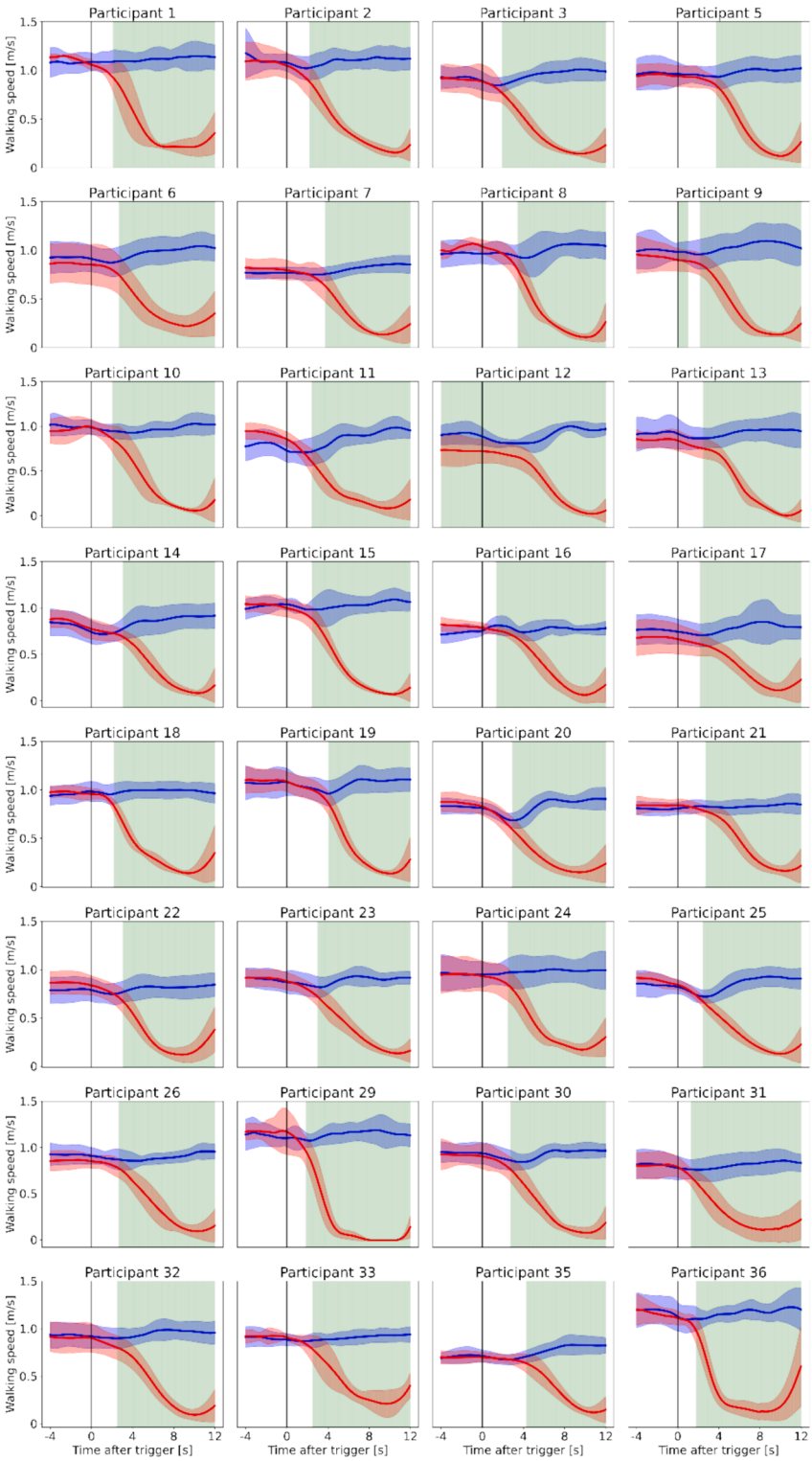


Fig. 4. Individual walking speed when crossing in front of vehicle (blue) or not (red). Green areas show significant differences between crossing and waiting trials (running *t*-test with false discovery rate correction). Excluded participants (#4, #27, #28, #34) not shown, see text for exclusion reasons. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

car, they walked on average with a speed of (0.95 ± 0.13) m/s (3.4 km/h) on the Omnideck. This is small compared to traffic observations of crossing pedestrians in the real world, which typically state 1.2–1.6 m/s (4.3–5.8 km/h) (Montufar et al., 2007; Aghabayk et al., 2021; Jain et al., 2014; Goh et al., 2012).

We replicate the three-stage-road-crossing model of (Gorrini et al., 2018) (cf. Fig. 3(b)) from real-traffic also in an omnidirectional walking simulation environment: At first, during the *approaching phase* (I), pedestrians walk with an approximately constant velocity of 0.95 m/s regardless of whether they will later cross before the car or not. Next, during the *appraising phase* (II), walking speed begins to decrease as the pedestrians approach the end of the alley. Interestingly, this is also the case if pedestrians later cross in front of the vehicle. This might be a precautionary measure to limit the deceleration in case of a later stopping decision or because of increased mental demand, since during this phase, visual search, perception and decision-making are likely taking place. During this phase, if crossing, pedestrians on average walk with 0.92 m/s. Finally, during the *crossing phase* (III) pedestrians keep decelerating in the case of waiting (in the "waiting" trials participants cross after the vehicle to continue their task), but accelerate in the case of crossing. During this phase, the walking speed slightly increases to 0.96 m/s, presumably because pedestrians quickly want to escape the danger of the approaching car. The walking speed phases were delineated based on changes in speed rather than absolute distance or time. To reduce noise in the speed data, the same moving-average filter with a window size of 0.8 s as (Gorrini et al., 2018) was applied. Following the approach of (Gorrini et al., 2018), a difference signal was computed by subtracting the mean of the smoothed speed data. The phase boundaries were determined by analysing the derivative of the difference signal. The onset of sustained deceleration, referred to as start of phase (II), was identified as the first instance where the derivative remained negative for at least one second. The start of phase (III) was then defined as the local minimum in the difference signal that followed (II), marking the moment when the pedestrian reached their lowest walking speed before accelerating again. Accordingly, phase (I) was defined as the period before the onset of sustained deceleration. (II) spanned from the onset of deceleration to the local minimum, while (III) began immediately after this minimum, corresponding to the pedestrian's acceleration. The phase-specific speed measures were extracted from the original (actual) walking speed data using these segmentations. However, a one-way ANOVA did not reveal a statistically significant difference in walking speed among the three phases ($p > 0.05$). When pedestrians decided to not cross in front of the car, the yielding takes more time compared to behaviour observed in real-traffic (Goldhammer et al., 2014). We attribute this to the specific treadmill mechanism of the Omnideck that transports the participants back to the centre of the simulator if they stop walking. As a consequence, the deceleration time takes approximately twice as long in our simulator experiment, compared to the 3.1 s that (Goldhammer et al., 2014) observed in real-traffic. Note that this issue is specific to substantial deceleration and should not (or far less) affect all of the other phases and velocity changes.

To investigate differences in walking speed between trials where participants crossed before vs. after the car, we averaged individual trials for these two categories for each participant (see Fig. 4). Three participants were excluded from all analyses, as the lack of one outcome (crossing/waiting) renders a meaningful analysis or prediction impossible: #28 and #34 never crossed before the car and #27 never crossed after the car, proving that it was basically possible to cross before every car. Additionally, #4 needed to be excluded for all further analyses in this manuscript due to implausibly slow walking speed in more than 20% of trials, which was likely caused by a technical error. Raw walking speed data was smoothed using the locally weighted scatter-plot smoothing (LOWESS) method (Cleveland, 1979) with $frac = 0.01$ smoothing parameter and no residual-based re-weightings ($it = 0$) to slightly reduce noise while closely following the overall trend of the data. To filter out trials where the walking speed in approaching the road was either initially too slow or implausibly high throughout the trial we used filters of 0.3 m/s at 4 s before the trigger and 2.0 m/s during the entire trial. The extreme values for the walking speeds were not caused by natural human behaviour but rather by rare technical errors, such as failures in tracking the head-mounted display and the resulting incorrect control of treadmill speed. Such technical errors were rare and affected only 15 of the 1,152 trials (32 included participants x 36 trials with vehicle), leaving 1,137 trials that were included in the walking-speed analysis. Crucially, no valid pedestrian behaviour was excluded based on these criteria: Speeds exceeding 2.0 m/s (7.2 km/h) were physically impossible on our Omnideck configuration and an initial speed below 0.3 m/s (1.1 km/h) would indicate non-compliance with our instructions to approach the road "*as you would in real life – without running but also without unnecessary delay*" (cf. Subsection 2.4).

We performed an independent *t*-test at each sampling point to compare walking speeds between the crossing and waiting conditions. To account for multiple comparisons, we – per participant – adjusted the alpha level to correspond to an expected false discovery rate (FDR) of 0.05 according to the procedure by (Benjamini and Hochberg, 1995). We asserted a significant walking speed difference between crossing and waiting trials for each sample point which fell below this corrected alpha level. The results revealed all time intervals (coloured green in Fig. 4), where the walking speed in crossing trials was significantly higher than in waiting trials. With the exception of participant #12 (who already walked significantly slower in waiting trials even before the trigger), did the two curves split significantly only after pedestrians passed the trigger and the car had been seen. The average time, where the participants significantly slowed down in the waiting trials was (2.52 ± 0.85) s after the trigger (excluding participant #12).

3.3. Head orientation

Importantly, the head yaw data did not indicate any reason for excluding any trial (cf. Fig. A.4). Thus, apart from the four excluded participants, no additional exclusions were necessary. This leaves all 1,152 trials (32 included participants x 36 trials with vehicle) for all head yaw-related analyses. When walking through the alley the head of the pedestrians is directed straight ahead in direction of walking. When approaching the end of the alley, we find that participants on average first turned their head to the left side by about 60° to check for approaching vehicles. Since cars were driving on the right side of the road in our experiment (as is usual in Germany), it is a good and usual strategy to look to the left side first (i.e., the approaching direction in the nearest lane). However, we do find that

some participants consistently looked to the right side first (e.g., participants #02 and #14 in Fig. A.4). The initial head-turn corresponds to the first peak in Fig. 6, revealing a first maximum of absolute head-turning frequency of about 40°/s around the time of the trigger, shortly before the pedestrians leave the alley. Thereby, the head is already pre-orientating to either side starting 2 s before reaching the trigger (~ 7.5 m before the kerb), presumably to detect the vehicle faster once the view to the road opens up. Between 0 and 1 s after the trigger (~ 5 m before the kerb), the mean absolute head-turning frequency is decreasing in all conditions (cf. Fig. 6), presumably because visual search for the car in the pre-decided direction and distance as well as speed/TTA estimation (Theisen et al., 2024) take place.

If participants correctly pre-oriented their head to the side where the car was approaching from, in the case of crossing, there is a phase of increased head movement starting from 1 s after the trigger (~ 4.5 m before the kerb) and peaking at 3–4 s after the trigger (~ 2 m before the kerb) at around 80–100°/s to make sure that road is safe to cross. The participants who did not cross before the car only reached approximately half of that head-turning velocity, which was also initiated 0.5–1 s later compared to the crossing cases. We think that this difference in absolute head-turning frequency indicates that the decision to cross has already been made before looking to the other side (at around 2–3 s after the trigger or ~ 3 m before the kerb), so this early and fast look to the other side (or lack that of) reveals an intention to (not) cross the road. From Fig. 5 it can be additionally observed that after the head-turn to the non-car side at around 2–3 s after the trigger, there is another head-turn back to the approaching car, presumably to check how fast the vehicle has been approaching and whether the original crossing decision is still appropriate. This so called 'last second check' (Hassan et al., 2005; Lyu et al., 2024; Yang et al., 2024) could also be a desire to make eye-contact with the driver to establish some form of communication, despite in our study it was stated that the cars would not react in any way to the participants' behaviour.

If participants do not find a car with their initial head-turn (i.e., when it was approaching from the other side instead), they immediately turn their head to the other side by approximately 120° from around 1 to 2 s after the trigger (~ 4 m before the kerb), searching for the car there. This corresponds with a second, more prominent maximum in absolute head-turning frequency of over 120°/s (see Fig. 6 right top) independent of whether the pedestrian later crossed the road in front of the car or not. In the following

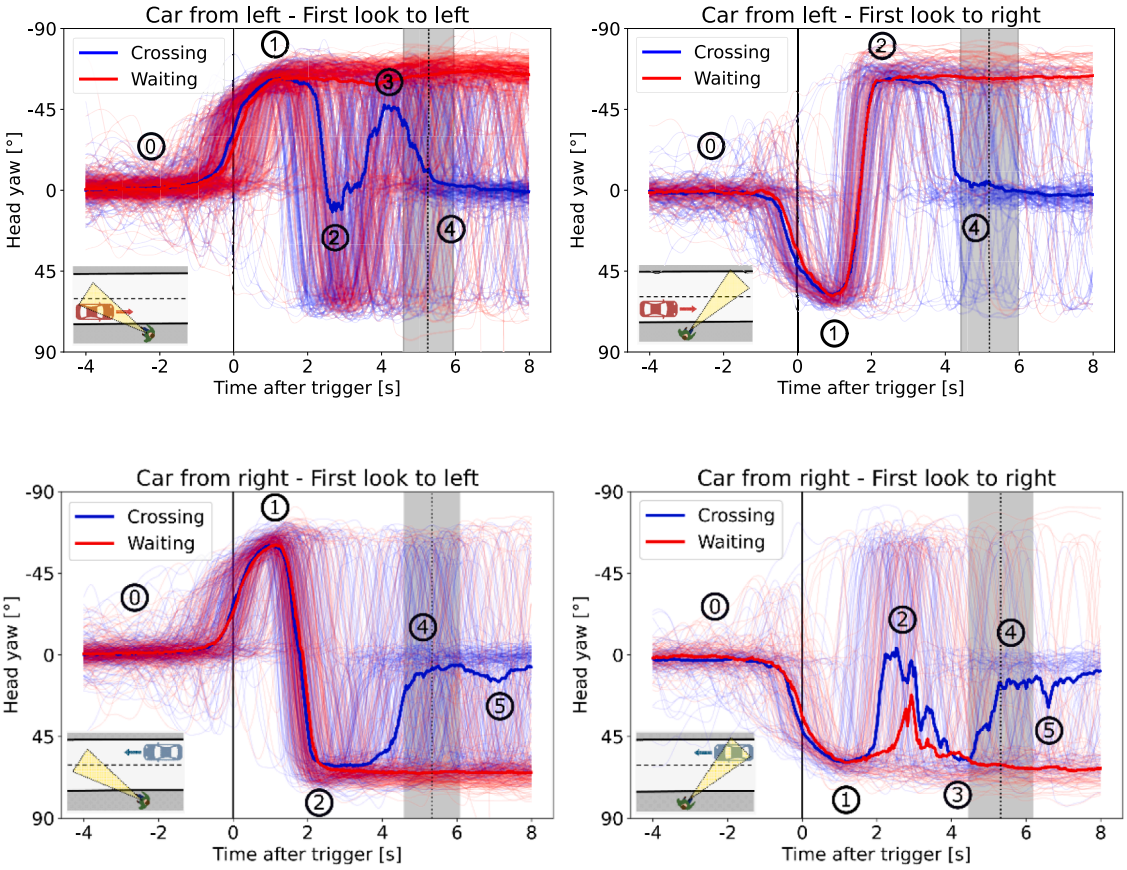


Fig. 5. Aggregated head yaw over all trials and all participants separated after crossing decision and first head turn direction. Negative values indicate a rotation to the left and positive values indicate a rotation to the right. Thick lines show median values per time frame. Dotted vertical line shows timing of crossing the kerb and shaded area shows SD. (0) Look straight while approaching the kerb, (1) Pre-orientation to correct (where the car is approaching from) or incorrect (where the car is not approaching from) side, (2) Check to other side, (3) Check again to approaching car ('last-second-check', (4) Look straight to cross the first lane, (5) Check again to approaching car in second lane and cross second lane.

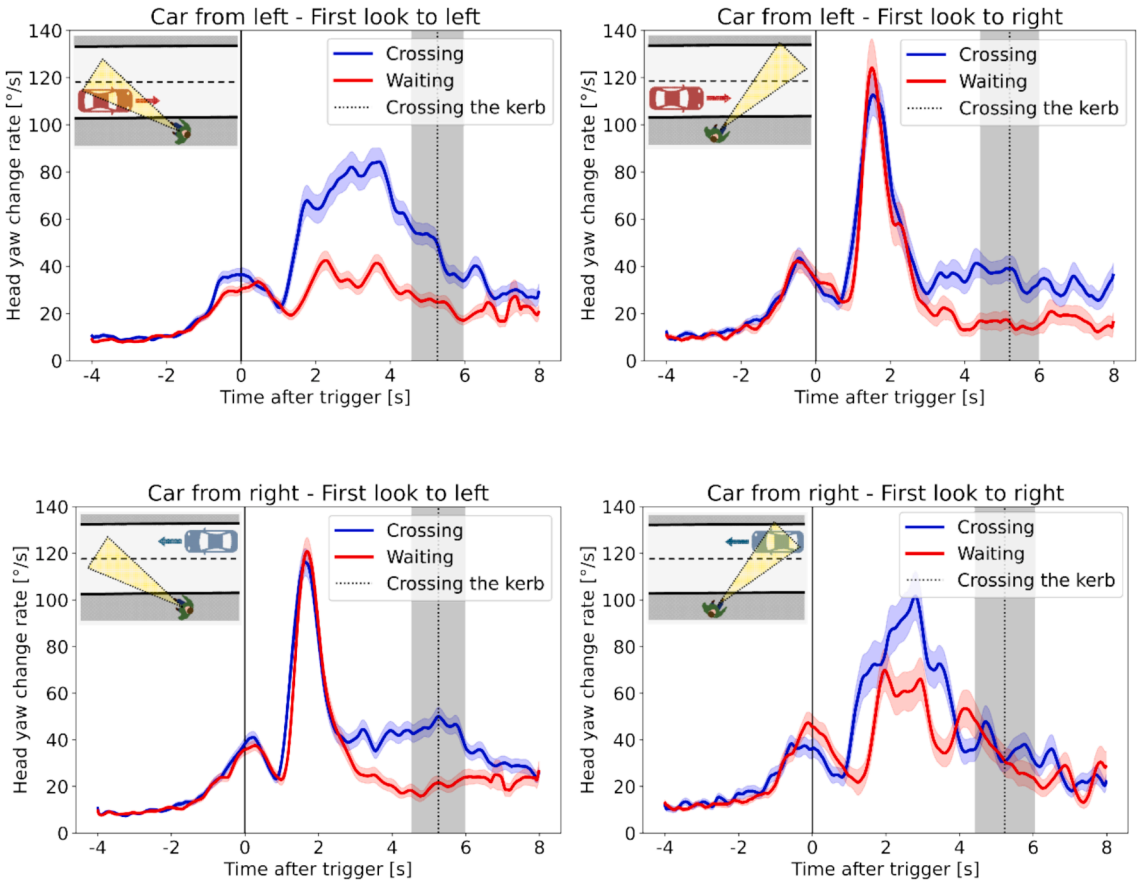


Fig. 6. Aggregated absolute head-turning frequency (°/s) over all trials and all participants separated after crossing decision and first head turn direction. Note that 'head-turning frequency' uses the absolute value of the turn rate; that is - for example − 60°/s over one second does not necessarily mean that the head turns by 60 degrees, but can also be a movement back and forth that results in no net turn. Blue/red solid lines show mean values and shaded areas show standard error of the mean (SEM). Black dotted lines show mean time when pedestrians crossed the kerb in the four cases, shaded areas show SD. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

second, the absolute head-turning frequency drops back down and splits up to an approximately constant frequency of 40°/s (when crossing) or 20°/s (when waiting), which are maintained during the next seconds. This is in accordance with Fig. 5 where it can be observed that when waiting, the pedestrians mostly kept looking at the approaching car, whereas in crossing cases, the head is turned back to the middle at the 4–5 s mark (∼ 1 m before the kerb) probably to focus on planning the steps to cross the road. This is again proceeded by a 'last-second-check' at around 7 s after the trigger (∼ 1.5 m after the kerb) towards the approaching car, if it drives in the second lane.

To summarise, looking away from the approaching car – or generally having more head-turning – after it had been initially detected before, could be an indication that the road will likely be crossed by a pedestrian. Importantly, the difference in head yaw between crossing and waiting cases is already showing between 1–2 s after the trigger (or ∼ 4 m before the kerb, see Fig. 6 left top), whereas the overt execution of the crossing decision is only visible around 5.5 s after the trigger, when pedestrians crossed the kerb. This suggests that head movements could be an early indicator for pedestrians' crossing intentions. An AV that detects such head movements could, for example, use this signal to inform its decision on slowing or stopping.

To test this, regarding the head orientation we performed a similar comparison between crossing and waiting trials as with the walking speed. Again, we smoothed the raw data using LOWESS, but due to the higher dynamic of the head yaw we smoothed with a fewer proportion of the data ($frac = 0.001$). Note that the filtering is not causal, but the potential error is bounded by the filter width, therefore negligible and identical for all such analyses. Similar to (Hamaoka et al., 2013; Lyu et al., 2024; Yang et al., 2024) we calculated the absolute value of the derivative of the head yaw to measure how fast the head is turning per time interval. According to the authors this head-turning frequency is a measure for the demand of information acquisition for the dynamic cues (e.g., distance, time-to-arrival) of the approaching vehicle. After first smoothing the raw signal to calculate the derivative on a cleaner signal with less high-frequency noise, we smooth a second time with LOWESS ($frac = 0.005$) since the differentiation amplified the noise. Next we performed pairwise t-tests with FDR correction (expected FDR of 0.05) at each time point between the crossing and waiting head-turning frequencies. The results revealed all time intervals (coloured green in Fig. 7), where the head-turning frequency in crossing

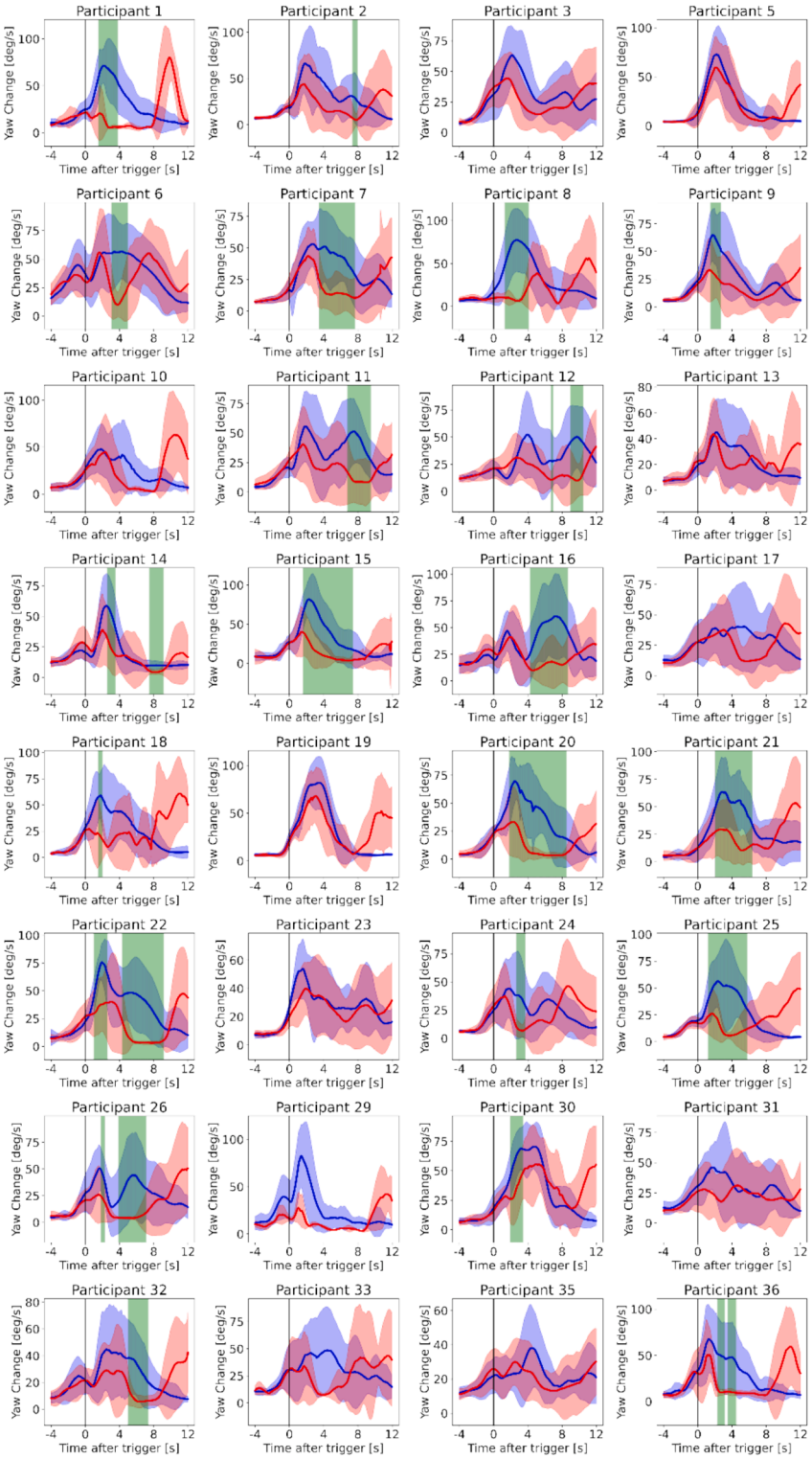


Fig. 7. Individual head-turning frequency when crossing in front of vehicle (blue) or not (red). Green areas show significant differences between crossing and waiting trials (running *t*-test with false discovery rate correction). Excluded participants (#4, #27, #28, #34) not shown, see text for exclusion reasons. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

trials was significantly higher than in waiting trials. However, we have noticed that the results are somewhat sensitive to the choice of smoothing filter strength *frac*. Specifically, while the overall trend remains consistent, the exact statistical significance varies across different smoothing configurations. This should be taken into account when interpreting the results (see limitations in Discussion section).

3.4. Trial-by-trial crossing prediction

While the average data suggests that head orientation could prove an early useful signal for decoding/predicting the crossing intention, the question remains whether this is also useful on a trial-by-trial basis. To this end, we compared the predictive power of the trajectory and the head orientation by performing a receiver operating characteristic (ROC) curve analysis for each participant. We analysed the mean walking speed and mean absolute head-turning frequency for crossing and waiting trials over different lengths of aggregated intervals (1 to 8 s) starting from the trigger point to assess how predictive power develops as more information becomes available. We assume that a crossing pedestrian is approaching the kerb with a higher speed and has a higher head-turning frequency due to a more intensive scanning process of the environment for approaching vehicles (e.g., for the purpose of safety assurance). Next we compared our prediction to the measured ground truth of crossing vs. waiting for each trial. Finally, we computed the area under the receiver operating characteristic curves (AUROC) per participant and calculate the mean values (see Fig. 8 for aggregated time intervals). To complement this, we conducted an additional analysis based on sequential 1-second intervals, which is presented in the Appendix (Figure A2). This provides a phase-specific examination of pedestrian decision-making, allowing for a more detailed assessment of decision dynamics over time. The analyses showed that already at 0–1 s after the trigger (~ 5 m before the kerb) – which is the earliest possible time when participants could detect a vehicle – both walking speed ($AUROC_{walk,0-1s} = 0.53 \pm 0.12$) and head-turning frequency ($AUROC_{head,0-1s} = 0.57 \pm 0.13$) had mean values above chance at correctly predicting the decision outcome. Importantly, in the early phase of the crossing (i.e., between 0–3 s after the trigger or ~ 5.5–2.5 m before the kerb) the head-turning frequency ($AUROC_{head,0-3s} = 0.71 \pm 0.15$) is a significantly stronger predictor for the crossing outcome than the walking speed ($AUROC_{walk,0-3s} = 0.64 \pm 0.13$). This is primarily caused by a more intense safety assurance process (i.e., the fast reorientation to the other side after having decided to cross) between 1–2 s (~ 4 m before the kerb), where the mean AUROC boxplot for head-turning frequency ($AUROC_{head,1-2s} = 0.67 \pm 0.14$) significantly outperforms the walking speed data ($AUROC_{walk,1-2s} = 0.60 \pm 0.11$, cf. Fig.A.1). From 3-4 s onwards (closer to the kerb than ~ 2 m), the walking speed is significantly outperforming our head-based predictor (presumably because pedestrians start decelerating when intending to stop, cf. Fig. 4) and finally becomes a perfect predictor at 5–6 s (around the time the kerb is crossed, $AUROC_{walk,5-6s} = 0.99 \pm 0.01$), whereas the head-based predictor converges, with the predictive power at a lower level ($AUROC_{head,5-6s} = 0.77 \pm 0.12$).

4. Discussion

We measured walking speed and head yaw of 36 pedestrians during road crossings (including the approach to the kerb) in a 360° treadmill walking simulator combined with VR. We find that in the first seconds after pedestrians notice the car, the absolute head-turning frequency is a stronger predictor for the crossing decision than the walking speed as measured by the AUROC. As time passes and pedestrians' decisions manifest themselves in either accelerating or decelerating, the walking speed becomes more useful than the head yaw and reaches an almost perfect predictive power of 0.99 at 5–6 s after the trigger, around the time the kerb is crossed. More specifically, decisions to cross seem to go hand in hand with a higher absolute head-turning frequency during the time after the car had initially been detected. So, while it is inevitable for an AV to track pedestrians' trajectories to reach a final decision on yielding to them, measuring head orientation will allow an earlier prediction, potentially enabling a more predictive, smoother and more human-like driving style.

The measured pedestrian behaviour is consistent with previous studies both regarding the walking speed, as well as the head-turning behaviour. Particularly, we replicate the three-phase-crossing-model of (Gorrini et al., 2018) and an increased head-turning during crossing trials (Kloeden et al., 2014; Lyu et al., 2024). More specifically, the measured pedestrians' head-turning behaviour in road crossings fits well to previous studies. Based on our data, we describe the following refined phases, which we have also summarised in a flowchart, see Figure A8:

(1) **Approaching**: Head is oriented straight in the direction of walking (Hassan et al., 2005)[1] and the walking speed is approximately constant (Gorrini et al., 2018).
(2) **Pre-orientation**: Head is turned towards one side (if the pre-orientation was towards the side without vehicle, this is corrected) (Geruschat et al., 2003; Yang et al., 2024); the walking speed begins to decrease.
(3) **Monitoring and decision**: Head remains aligned with the approaching vehicle (for the perception of dynamic cues); the crossing decision is made and walking speed reaches a minimum (Gorrini et al., 2018).

[1] The initial forward gaze in our study may have been reinforced by the guided alley, as there was no possibility to scan for traffic inside the alley. This design choice was made to synchronise head movements at the trigger point. However, forward-looking behaviour when approaching the kerb perpendicularly is generally expected in structured environments, as pedestrians tend to orient their head in the direction of walking before shifting attention to traffic. While in more open environments, head movements may begin earlier, visibility restrictions – such as parked cars, buildings, or street furniture – are common, meaning that scanning patterns may still be somewhat delayed.

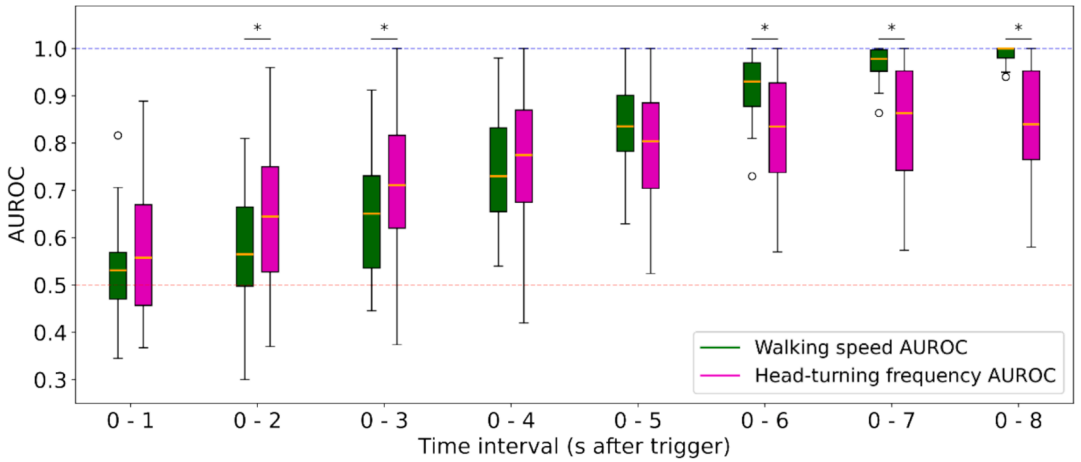


Fig. 8. Mean AUROC boxplot for walking speed (green) and head-turning frequency (magenta) data using different aggregated time intervals of data after the trigger. AUROC values of 1.0 mean perfect decoding (blue dashed line), 0.5 means chance (red dashed line). The boxes extend from the first to the third quartile of the data (forming the inter-quartile range [IQR]), with an orange line at the median. The whiskers extend from the box to the farthest data point lying within 1.5xIQR from the box. Outlier points are those past the end of the whiskers. Asterisks (*) indicate statistically significant differences between conditions ($p < 0.05$), as determined by pairwise t-tests. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(4) **Safety assurance** (if intended to cross): Head-turning frequency increases starting shortly before and peaking shortly after crossing initiation (each side is looked at at least once) (Whitebread and Neilson, 2001; Hassan et al., 2005; Hamaoka et al., 2013; Yang et al., 2024; Lyu et al., 2024); the walking speed begins to increase.
(5) **Finalisation of crossing decision**: Head is directed toward the approaching vehicle for a 'last-second-check' before crossing the kerb, regardless of which lane the vehicle is in; if a vehicle is in the second lane, an additional, less pronounced check occurs before entering that lane (Geruschat et al., 2003; Hamaoka et al., 2013); walking speed increases further.
(6) **Execution**: Pedestrians cross the road and look in the direction of walking; after the road the walking speed decreases and stabilises (Gorrini et al., 2018).

This replicates and refines earlier findings for general boundary conditions, in particular by including the approach to the road (compared to previous studies starting from standstill at the kerb), as well as two-lane roads where the pedestrian is not previously aware of the direction from which a vehicle might approach and various combinations of vehicle speed and time gap. Additionally, we provide evidence that head movements away from the vehicle (and generally more head-turning) after having detected it likely announce a crossing decision while an ongoing fixation of the approaching vehicle likely announce a waiting decision (Kloeden et al., 2014).

These results have some implications: Firstly, we suggest that head orientation (specifically absolute head-turning frequency) could be used as a feature for predicting pedestrians' crossing intentions during the initial phase of visual search, perception and decision-making while other trajectory-based features (e.g., walking speed) should be emphasised later, when decisions are operationalised. This time-dependent prediction model will likely outperform models that solely rely on pedestrians' trajectories. This is especially relevant, since (Rehder et al., 2014) and (Flohr et al., 2015) have shown the possibility for real-time recognition of head orientation from camera images. In this regard, we think that the head information could be used by automated vehicles to adapt their driving strategies, for example, by increasing the probability to yield if a pedestrian is looking to the opposite direction (or generally is showing more head-turning) after having looked in the direction of the AV. Hereby, we hope that our results are of interest for engineers who already have used head information in pedestrian prediction (Kooij et al., 2014; Schulz and Stiefelhagen, 2015; Kloeden et al., 2014). However, our results come with some limitations: One limitation of our analysis is the necessity to choose the smoothing parameters, in particular the filter strength (*frac*) for the LOWESS filter. We chose the *frac* parameter based on visual inspection of raw data (i.e., prior to any statistical testing) as a compromise between filtering out high-frequency noise and preserving the overall temporal dynamics. After fixing the parameter and conducting the analysis as reported, we tested the robustness of our findings by varying the *frac* parameter over two orders of magnitude and recomputed the mean predictive power (AUROC) of the head-turning frequency. The key result pattern was robust to this variation of the *frac* parameter: during the approach to the kerb, AUROCs based on head-turning frequency remained consistently higher than AUROCs based on walking speed, underlining the relevance of head-turning frequency for distinguishing crossing from waiting trials. However, in some cases, the difference did not reach statistical significance at the conventional alpha level of 5%, highlighting that smoothing choices necessarily have some quantitative effects. Future work could therefore specifically explore alternative filtering methods and parameters to further strengthen the validity and robustness of the findings. Nonetheless, the qualitative pattern of when head-turning frequency outperforms walking speed is remarkably robust to the specific smoothing choices.

Since our data was obtained using a HMD combined with VR, results might lack comparability to real-world traffic. The question to

what extent pedestrians' gap acceptance differs in virtual environments compared to the real world is still debated. (Feldstein and Dyszak, 2020) reported that pedestrians accepted 26% smaller minimal time-gaps in VR compared to a real test-track study. Oppositely, (Schneider et al., 2022) demonstrated that participants have a reduced crossing probability in virtual environments, especially if HMDs were used. Although we used different time gaps and velocities between trials, in our study, vehicles maintained a constant speed without responding to pedestrians, as there was no designated pedestrian crossing requiring them to yield. In reality, human drivers frequently adjust their behaviour in response to pedestrians, which in turn influences crossing decisions. To focus solely on pedestrian-initiated crossing behaviour, we opted for a non-responsive vehicle behaviour to maintain a controlled experimental environment. While this approach ensures internal validity, it does not fully capture the interactive dynamics of pedestrian-vehicle interactions. A necessary next step will be to integrate adaptive vehicle behaviour to test our predictive model under more realistic conditions and explore how pedestrians adjust their decisions in response to different vehicle behaviour, for example acceleration (Wessels et al., 2023). In addition, we did not provide auditory cues, which is known to influence pedestrian gap acceptance (Wessels et al., 2023).

In particular, the use of a treadmill-based walking simulator requires a special discussion of results. Firstly, the mean measured walking speed during crossing trials of 3.4 km/h was small compared to real-world observations (Montufar et al., 2007; Aghabayk et al., 2021; Jain et al., 2014; Goh et al., 2012). Secondly, step size is known to be reduced in VR (Hollman et al., 2006), but even more so when VR is combined with treadmill simulators in general (Sloot et al., 2014), with the Omnideck being no exception (Temme et al., 2024). Thirdly, the stopping time of pedestrians on the Omnideck is artificially increased, as it uses a return mechanism to transport the user back to the static platform in the middle, as this enables quick changes of direction. Furthermore, it is important to note that HMDs have a small field-of-view (FOV = 110° for the used Vive Pro Eye as specified by HTC cooperation), which is approximately only half (Strasburger, 2020) of what human eyes are capable of. Experiments showed that participants move their heads more (displacement) the smaller the FOV (Wells and Venturino, 1990), as peripheral vision is more limited. Counterintuitively, participants also moved their heads slower the smaller the FOV was (Wells and Venturino, 1990; Venturino and Wells, 1990), which the authors attributed to a combination of less certainty for target positions and a fear for overshooting in the used visual search task. For the context of our VR road crossing task this means that pedestrians' head-turning frequency might even be bigger in real-traffic than as measured here in VR. Intentionally, we used an artificially empty traffic scenario without other road users or vegetation to limit the influence of distractions. However, this limits the real-world applicability of our results. In real-traffic pedestrians' attention (and consequently head movements) is influenced by salient objects (especially in children (Tapiro et al., 2020)), for example, advertisements or other pedestrians, limiting the validity of head movements as an indicator of the pedestrian's future whereabouts.

Lastly, the unequal gender distribution in our sample may impact the generalisability of our findings. Observational studies from Canada (Montufar et al., 2007), India (Jain et al., 2014), Iran (Aghabayk et al., 2021), and Malaysia (Goh et al., 2012) consistently report that male pedestrians walk and cross faster than females. Additionally, (Holland and Hill, 2010) found that men take greater risks, accepting smaller safety margins and engaging in more unsafe crossings than women. (Tom and Granié, 2011) further observed gender differences in gaze patterns, with women focusing more on other pedestrians while men rather concentrated on vehicles before and during crossing. Given that walking speed, head movements, and risk perception are key factors in our study, an unequal gender distribution could impact the predictive strength of head movements for crossing decisions. Future research should aim for a balanced gender sample to enhance the robustness and generalisability of the findings.

To summarise, we advance previous pedestrian crossing studies by analysing pedestrians' walking speed and head orientation including the approach to the kerb, a two-lane road and traffic from both sides. We provide evidence for the predictive ability of pedestrians' head turns regarding the crossing decision by comparing it to the walking speed in a ROC analysis. Thereby, we aim to close the gap between psychology and engineering by providing evidence for the absolute head-turning frequency as a possible feature in prediction algorithms. Future studies should investigate pedestrians' head orientation in more naturalistic settings including multi-vehicle and multi-pedestrian environments. Our setting can readily be extended to such scenarios. Thereby, we can enable automated driving systems to help reduce crashes with pedestrians by predicting their future behaviour, which this study aims to advance in a meaningful way.

CRediT authorship contribution statement

Max Theisen: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Melina Bergen:** Writing – review & editing, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Wolfgang Einhäuser:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Caroline Schießl:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

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.

Acknowledgments

Max Theisen, Melina Bergen, and Caroline Schießl received support from the Japanese-German Research Collaboration on Human Factors in Connected and Automated Driving (CADJapanGermany) supported by the German Federal Ministry of Education and Research (funding code ID 16ES1038). Wolfgang Einhäuser received support from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project ID 416228727 - SFB 1410: Hybrid Societies: Humans Interacting with Embodied Technologies.

Appendix

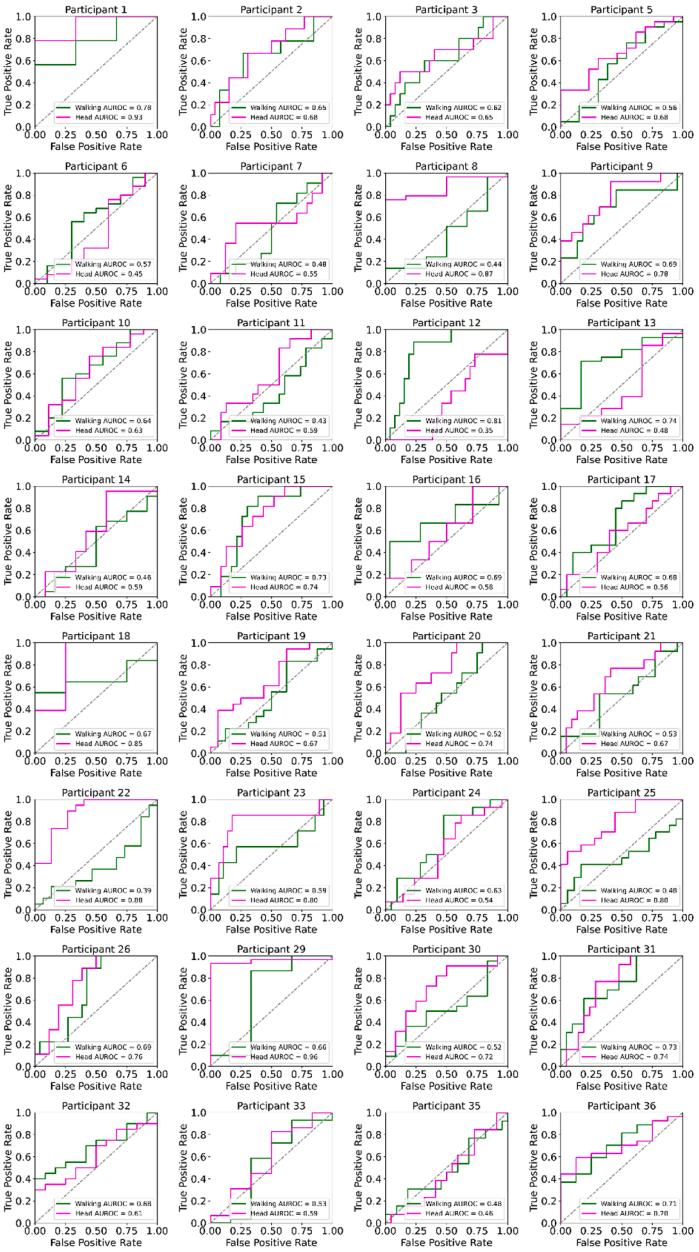


Fig. A1. Exemplary ROC curves per person using walking speed (green) and head-turning frequency (magenta) between **1 - 2 s** after the trigger as predictors for the crossing decisions. The area under the ROC curve (AUROC) indicates the classification performance, with higher values reflecting better predictive accuracy. Diagonal dashed lines represent chance-level performance. Excluded participants (#4, #27, #28, #34) not shown, see text for exclusion reasons.

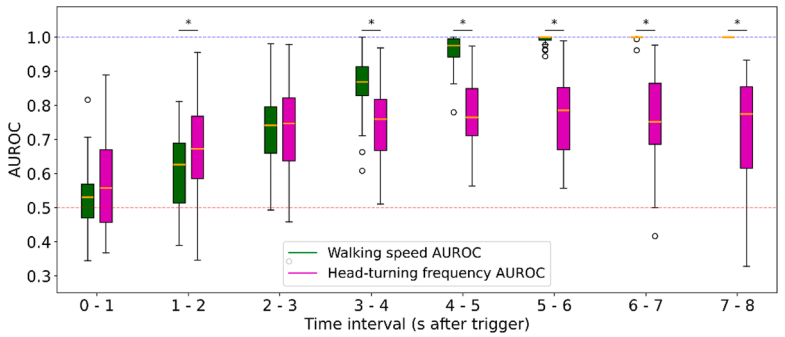


Fig. A2. Mean AUROC boxplot for walking speed (green) and head-turning frequency (magenta) data using different 1 s time intervals of data after the trigger. AUROC values of 1.0 mean perfect decoding (blue dashed line), 0.5 means chance (red dashed line). The boxes extend from the first to the third quartile of the data (forming the inter-quartile range [IQR]), with an orange line at the median. The whiskers extend from the box to the farthest data point lying within 1.5xIQR from the box. Outlier points are those past the end of the whiskers. Asterisks (*) indicate statistically significant differences between conditions ($p < 0.05$), as determined by pairwise t-tests.

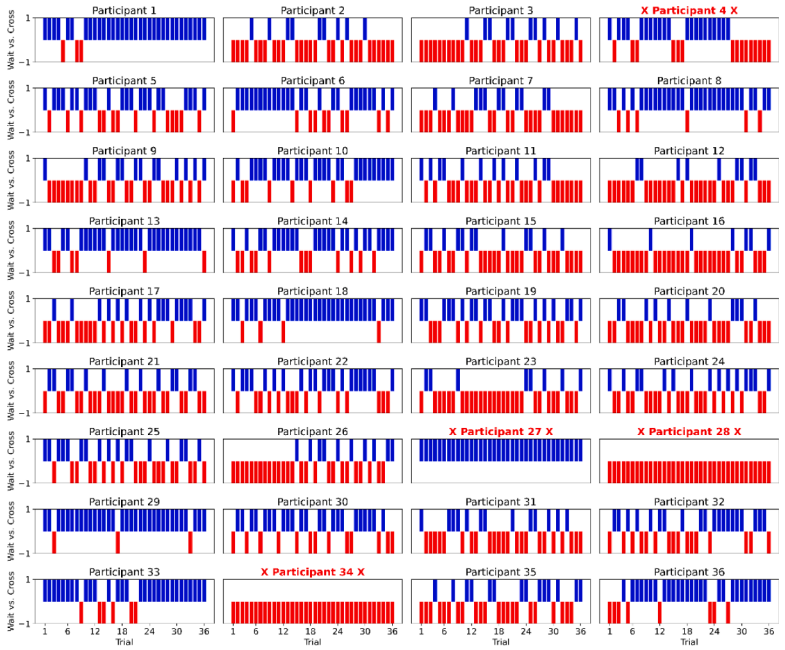


Fig. A3. Individual crossing histograms for all trials in the order of presentation. Blue bars show trials where participants crossed before the car, red bars show trials where participants crossed after the car. Participants marked with "X" in red are excluded from further analysis, see text for exclusion reasons.

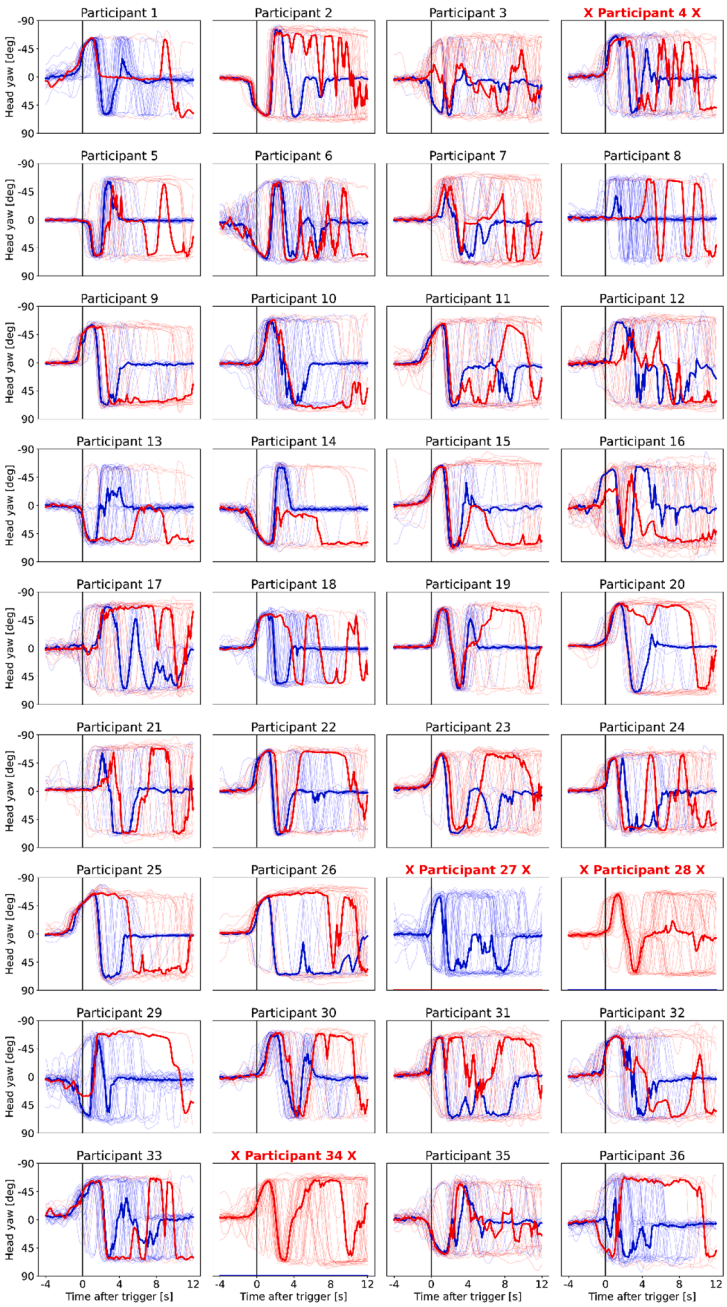


Fig. A4. Individual head yaw over time relative to the trigger, with negative values indicating a rotation to the left and positive values indicating a rotation to the right. Blue lines indicate trials where the pedestrian was crossing before the car, while red lines indicate trials where the pedestrian was crossing after the car. Thick lines show medians. Participants marked with "X" in red are excluded from further analysis, see text for exclusion reasons.

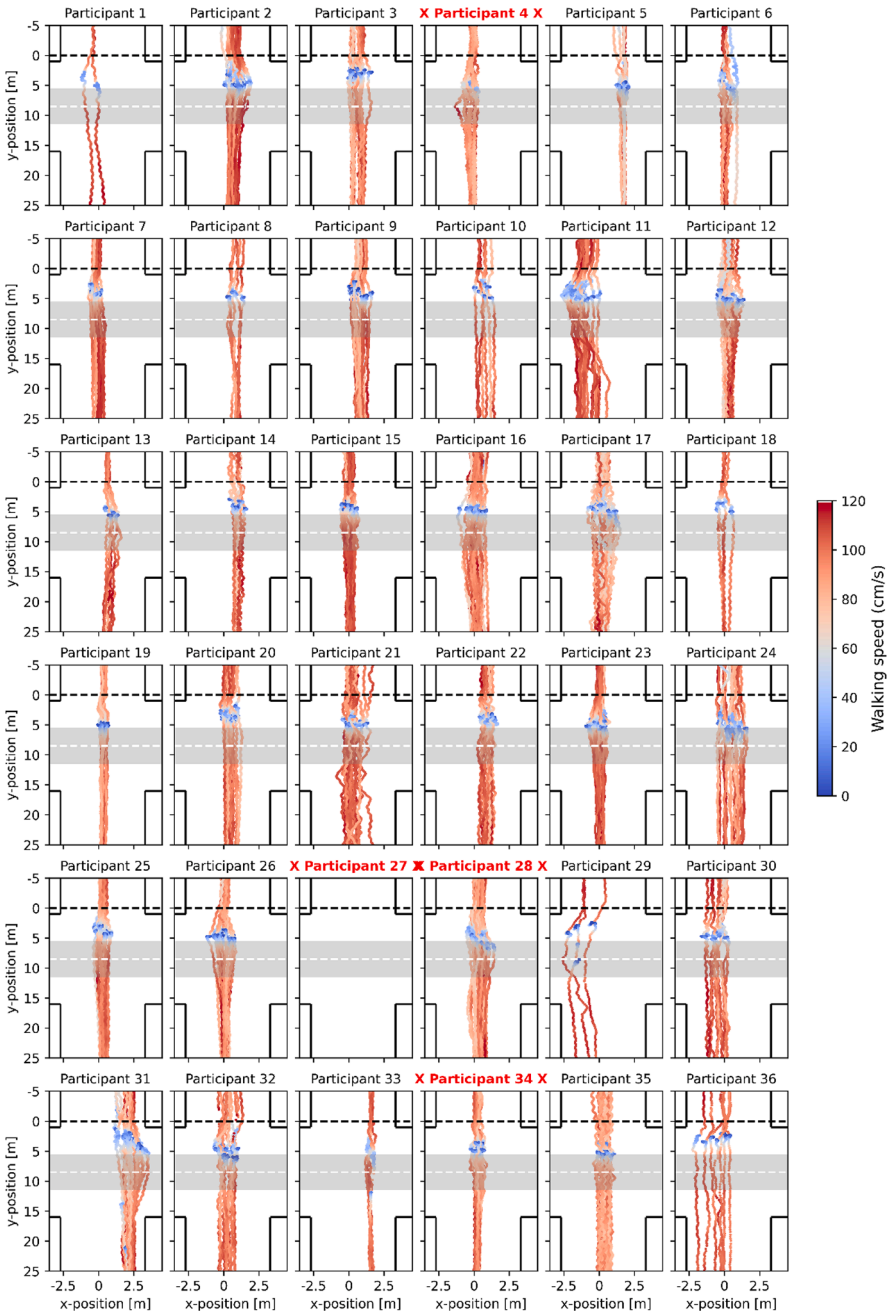


Fig. A5. Trajectories of pedestrians from top view in trials where the pedestrians yielded to the vehicle and crossed after it. Walking speeds (cm/s) are colour-coded. Pedestrians were walking from negative y-positions (top) to positive y-positions (bottom). Shaded grey regions show the road with the dashed white centre line. Dashed black lines show the (invisible) trigger and solid black lines show the outline of buildings. Participants marked with "X" in red were excluded from further analysis, see text for exclusion reasons. Due to image size, only every 20th data point is shown.

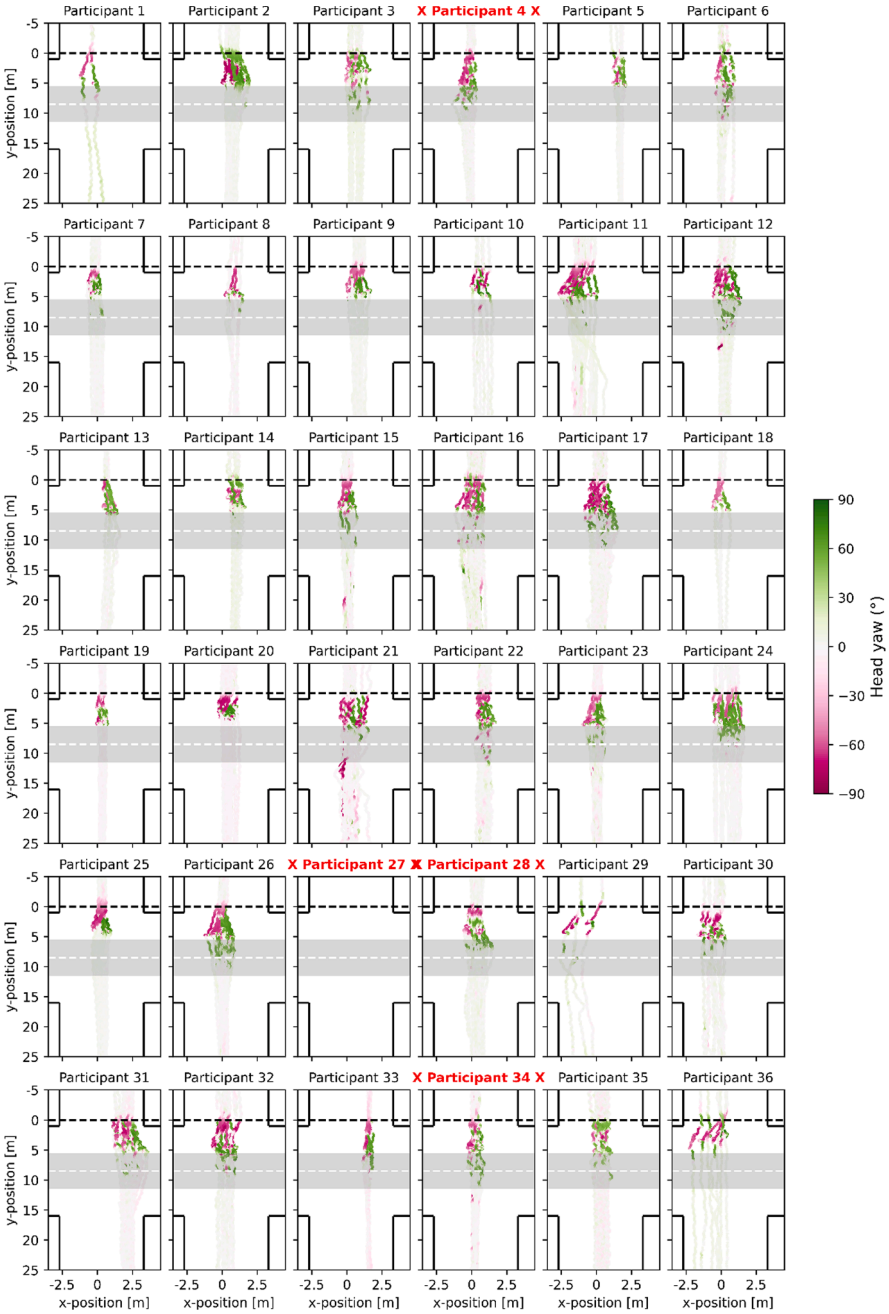


Fig. A6. Trajectories of pedestrians from top view in trials where the pedestrians yielded to the vehicle and crossed after it. Head yaw (deg) is colour-coded (negative values mean looking to left). Pedestrians were walking from negative y-positions (top) to positive y-positions (bottom). Shaded grey regions show the road with the dashed white centre line. Dashed black lines show the (invisible) trigger and solid black lines show the outline of buildings. Participants marked with "X" in red were excluded from further analysis, see text for exclusion reasons. Due to image size, only every 20th data point is shown.

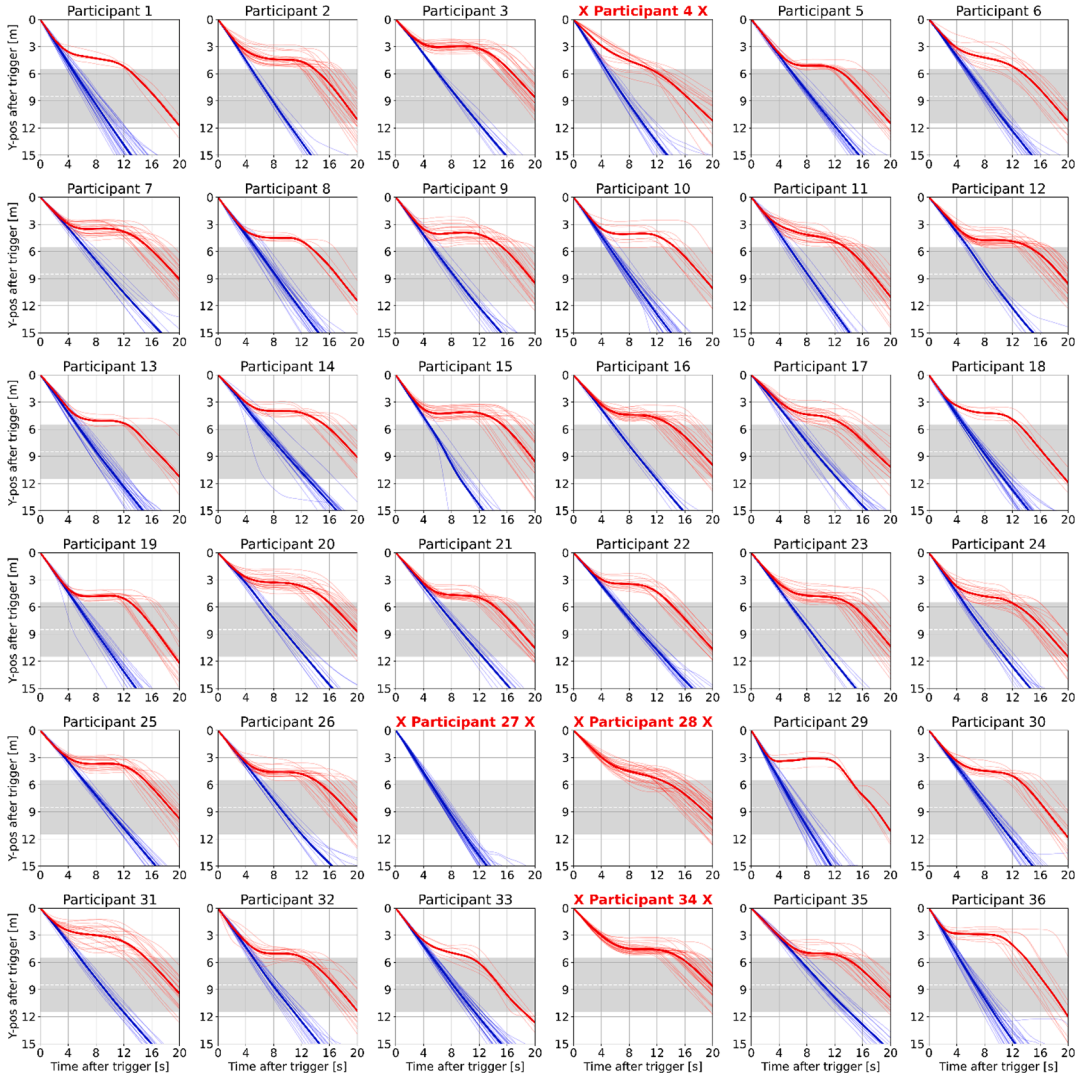


Fig. A7. Y-position of pedestrians over time relative to the trigger. Blue/red lines indicate trials where the pedestrian was crossing before/after the car. Thick lines show mean values. Shaded grey regions show the road with the dashed white centre line. Participants marked with "X" in red were excluded from further analysis, see text for exclusion reasons.

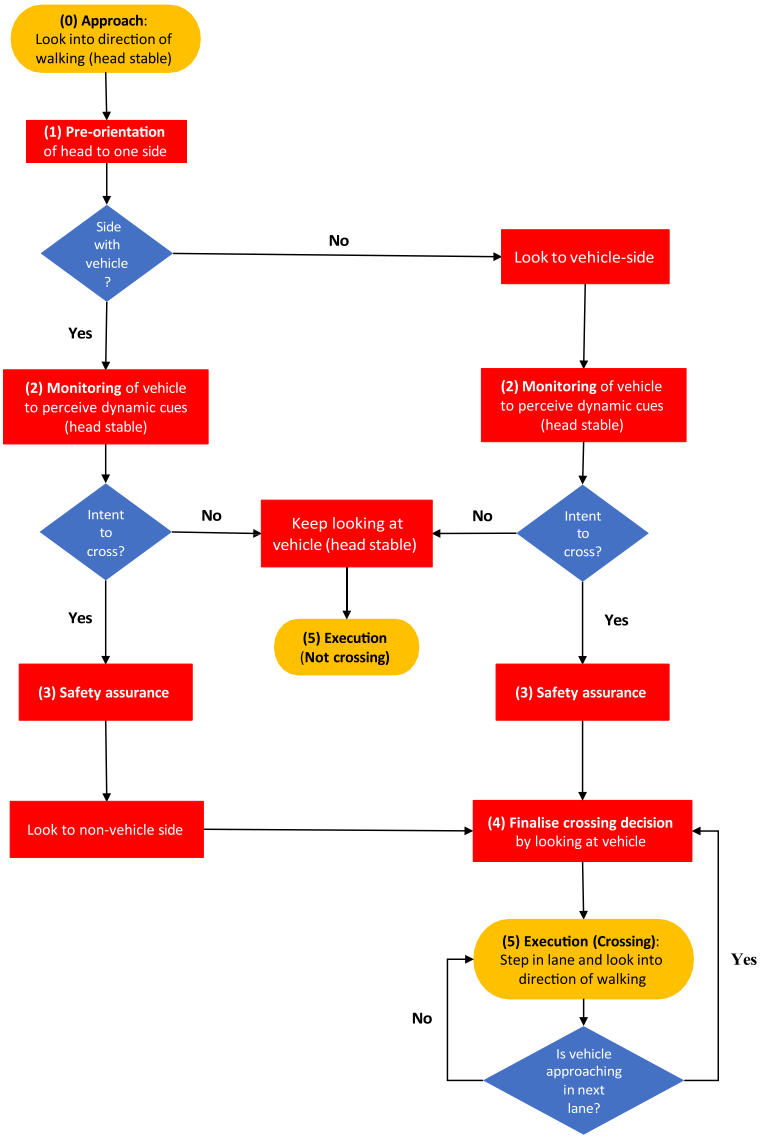


Fig. A8. Flowchart of the pedestrian crossing decision process regarding head movements.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trf.2025.05.015.

Data availability

Data will be made available on request.

References

Aghabayk, K., Esmailpour, J., Jafari, A., & Shiwakoti, N. (2021). Observational-based study to explore pedestrian crossing behaviors at signalized and unsignalized crosswalks. *Accident Analysis & Prevention, 151*, Article 105990.

AlAdawy, D., Glazer, M., Terwilliger, J., Schmidt, H., Domeyer, J., Mehler, B., Reimer, B., and Fridman, L. (2019). Eye contact between pedestrians and drivers. *arXiv preprint arXiv:1904.04188*.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300.

Brinks, H. and Bruins, M. (2016). Redesign of the omnideck platform: With respect to dfa and modularity.

Brunetti, A., Buongiorno, D., Trotta, G. F., & Bevilacqua, V. (2018). Computer vision and deep learning techniques for pedestrian detection and tracking: A survey. *Neurocomputing, 300*, 17–33.
Cao, J., Pang, Y., Xie, J., Khan, F. S., & Shao, L. (2021). From handcrafted to deep features for pedestrian detection: A survey. *IEEE transactions on pattern analysis and machine intelligence, 44*(9), 4913–4934.
Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association, 74*(368), 829–836.
Cohen, J., Dearnaley, E., & Hansel, C. (1955). The risk taken in crossing a road. *Journal of the Operational Research Society, 6*(3), 120–128.
Cox, D. and Snell, E. (1989). *Analysis of Binary Data, Second Edition*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis.
Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 886–893. Ieee.
Fagnant, D. J. and Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, 77:167– 181.
Feldstein, I. T. and Dyszak, G. N. (2020). Road crossing decisions in real and virtual environments: A comparative study on simulator validity. *Accident Analysis & Prevention*, 137:105356.6.
Flohr, F., Dumitru-Guzu, M., Kooij, J. F., & Gavrila, D. M. (2015). A probabilistic framework for joint pedestrian head and body orientation estimation. *IEEE Transactions on Intelligent Transportation Systems, 16*(4), 1872–1882.
Geruschat, D. R., Hassan, S. E., & Turano, K. A. (2003). Gaze behavior while crossing complex intersections. *Optometry and vision science, 80*(7), 515–528.
Goh, B. H., Subramaniam, K., Wai, Y. T., Mohamed, A. A., & Ali, A. (2012). Pedestrian crossing speed: The case of malaysia. *International Journal for traffic and transport engineering, 2*(4), 323–332.
Goldhammer, M., Hubert, A., Koehler, S., Zindler, K., Brunsmann, U., Doll, K., and Sick, B. (2014). Analysis on termination of pedestrians' gait at urban intersections. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1758–1763. IEEE.
Gorrini, A., Crociani, L., Vizzari, G., & Bandini, S. (2018). Observation results on pedestrian- vehicle interactions at non-signalized intersections towards simulation. *Transportation Research Part F Traffic Psychology and Behaviour, 59*, 269–285.
Guéguen, N., Meineri, S., & Eyssartier, C. (2015). A pedestrian's stare and drivers' stopping behavior: A field experiment at the pedestrian crossing. *Safety science, 75*, 87–89.
Hamaoka, H., Hagiwara, T., Tada, M., & Munehiro, K. (2013). A study on the behavior of pedestrians when confirming approach of right/left-turning vehicle while crossing a crosswalk. *Journal of the Eastern Asia Society for Transportation Studies, 10*, 2109–2122.
Hassan, S. E., Geruschat, D. R., & Turano, K. A. (2005). Head movements while crossing streets: Effect of vision impairment. *Optometry and vision science, 82*(1), 18–26.
Holland, C. and Hill, R. (2010). Gender differences in factors predicting unsafe crossing decisions in adult pedestrians across the lifespan: A simulation study. *Accident Analysis & Prevention*, 42(4):1097– 1106.
Hollman, J. H., Brey, R. H., Robb, R. A., Bang, T. J., & Kaufman, K. R. (2006). Spatiotemporal gait deviations in a virtual reality environment. *Gait & posture, 23*(4), 441–444.
Jain, A., Gupta, A., and Rastogi, R. (2014). Pedestrian crossing behaviour analysis at intersections. *International Journal for Traffic and Transport Engineering*, 4(1): 103–116.
Kalantarov, S., Riemer, R., & Oron-Gilad, T. (2018). Pedestrians' road crossing decisions and body parts' movements. *Transportation research part F: traffic psychology and behaviour, 53*, 155–171.
Keller, C. G. and Gavrila, D. M. (2013). Will the pedestrian cross? a study on pedestrian path prediction. *IEEE Transactions on Intelligent Transportation Systems*, 15(2): 494–506.
Kloeden, H., Brouwer, N., Ries, S., and Rasshofer, R. H. (2014). Potenzial der kopfposenerken- nung zur absichtsvorhersage von fußgängern im urbanen verkehr. In *FAS Workshop Fahrerassistenzsysteme, Walting, Germany*.
Koehler, S., Goldhammer, M., Bauer, S., Zecha, S., Doll, K., Brunsmann, U., & Dietmayer, K. (2013). Stationary detection of the pedestrian? s intention at intersections. *IEEE Intelligent Transportation Systems Magazine, 5*(4), 87–99.
Kooij, J. F. P., Schneider, N., Flohr, F., and Gavrila, D. M. (2014). Context-based pedestrian path prediction. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VI 13*, pages 618–633. Springer.
Kooijman, L., Happee, R., & de Winter, J. C. (2019). How do ehmis affect pedestrians' crossing behavior? a study using a head-mounted display combined with a motion suit. *Information, 10*(12), 386.
Korbmacher, R. and Tordeux, A. (2022). Review of pedestrian trajectory predic- tion methods: Comparing deep learning and knowledge-based approaches. *IEEE Transactions on Intelligent Transportation Systems*, 23(12):24126–24144.
Lee, Y. M., Madigan, R., Giles, O., Garach-Morcillo, L., Markkula, G., Fox, C., Camara, F., Roth- mueller, M., Vendelbo-Larsen, S. A., Rasmussen, P. H.,, et al. (2021). Road users rarely use explicit communication when interacting in today's traffic: Implications for automated vehicles. *Cognition, Technology & Work, 23*, 367–380.
Litman, T. (2020). *Autonomous vehicle implementation predictions: Implications for transport planning*. Victoria Transport Policy Institute: Technical report.
Lobjois, R. and Cavallo, V. (2007). Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task. *Accident analysis & prevention*, 39(5):934–943.
Lubbe, N., Wu, Y., & Jeppsson, H. (2022). Safe speeds: Fatality and injury risks of pedestrians, cyclists, motorcyclists, and car drivers impacting the front of another passenger car as a function of closing speed and age. *Traffic safety research, 2*, Article 000006.
Lyu, W., Mun Lee, Y., Uzondu, C., Madigan, R., Gonçalves, R. C., Garcia de Pedro, J., Romano, R., & Merat, N. (2024). A distributed simulation study to investigate pedestrians' road-crossing decisions and head movements in response to different vehicle kinematics in mixed traffic. *Transportation Research Part F: Traffic Psychology and Behaviour, 104*, 1–14.
Mínguez, R. Q., Alonso, I. P., Fernández-Llorca, D., & Sotelo, M. A. (2018). Pedestrian path, pose, and intention prediction through gaussian process dynamical models and pedestrian activity recognition. *IEEE Transactions on Intelligent Transportation Systems, 20*(5), 1803–1814.
Montufar, J., Arango, J., Porter, M., & Nakagawa, S. (2007). Pedestrians' normal walking speed and speed when crossing a street. *Transportation research record, 2002* (1), 90–97.
Novacek, T. and Jirina, M. (2022). Overview of controllers of user interface for virtual reality. *Presence (Camb.)*, 29:37–90.
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.
Pekkanen, J., Giles, O. T., Lee, Y. M., Madigan, R., Daimon, T., Merat, N., & Markkula, G. (2022). Variable-drift diffusion models of pedestrian road-crossing decisions. *Computational Brain & Behavior*, 1–21.
Petzoldt, T. (2014). On the relationship between pedestrian gap acceptance and time to arrival estimates. *Accid. Anal. Prev.,, 72*, 127–133.
Pipkorn, L., Noonan, T. Z., Domeyer, J., Mehler, B., Reimer, B., and Gershon, P. (2024). Naturalistic analysis of bidirectional gazing during vehicle-pedestrian encounters at road crossings. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 68, pages 960–964. SAGE Publications Sage CA: Los Angeles, CA.
Rasouli, A., Kotseruba, I., and Tsotsos, J. K. (2017a). Agreeing to cross: How drivers and pedestrians communicate. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pages 264–269. IEEE.
Rasouli, A., Kotseruba, I., & Tsotsos, J. K. (2017b). Understanding pedestrian behavior in complex traffic scenes. *IEEE Transactions on Intelligent Vehicles, 3*(1), 61–70.
Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779–788).
Rehder, E., Kloeden, H., and Stiller, C. (2014). Head detection and orientation estimation for pedestrian safety. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 2292–2297. IEEE.
Ren, Z., Jiang, X., & Wang, W. (2016). Analysis of the influence of pedestrians' eye contact on drivers' comfort boundary during the crossing conflict. *Procedia engineering, 137*, 399–406.

Ridel, D., Rehder, E., Lauer, M., Stiller, C., and Wolf, D. (2018). A literature review on the predic- tion of pedestrian behavior in urban scenarios. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 3105–3112. IEEE.

Schmidt, S. and Faerber, B. (2009). Pedestrians at the kerb–recognising the action intentions of humans. *Transportation research part F: traffic psychology and behaviour*, 12(4):300–310.

Schmidt, S., Färber, B., and Grassi, A. P. (2008). Geht er oder geht er nicht?-ein fas zur vorhersage von fußgängerabsichten. In *5. Workshop Fahrerassistenzsysteme*.

Schneider, S., Maruhn, P., Dang, N.-T., Pala, P., Cavallo, V., & Bengler, K. (2022). Pedestrian crossing decisions in virtual environments: Behavioral validity in caves and head-mounted displays. *Human factors, 64*(7), 1210–1226.

Schulz, A. T. and Stiefelhagen, R. (2015). A controlled interactive multiple model filter for combined pedestrian intention recognition and path prediction. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 173–178. IEEE.

Singh, S. (2015). *Critical reasons for crashes investigated in the national motor vehicle crash causation survey*. National Highway Traffic Safety Administration (NHTSA): Technical report.

Sloot, L., Van der Krogt, M., & Harlaar, J. (2014). Effects of adding a virtual reality environment to different modes of treadmill walking. *Gait & posture, 39*(3), 939–945.

Sobhani, A. and Farooq, B. (2018). Impact of smartphone distraction on pedestrians' crossing behaviour: An application of head-mounted immersive virtual reality. *Transportation research part F: traffic psychology and behaviour*, 58:228–241.

Strasburger, H. (2020). Seven myths on crowding and peripheral vision. *i-Perception*, 11(3):2041669520913052.

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. *Journal of safety research, 72*, 101–109.

Temme, G., Fischer, M., Bergen, M., Grone, K., Rehm, M., & Wegener, J. (2024). In *Development of a playful vr training sequence for a treadmill-based pedestrian simulator* (pp. 30–36). Strasbourg, France: Driving Simulation Association.

Theisen, M., Schießl, C., Einhäuser, W., & Markkula, G. (2024). Pedestrians' road-crossing de- cisions: Comparing different drift-diffusion models. *International Journal of Human-Computer Studies, 183*, Article 103200.

Tian, K., Markkula, G., Wei, C., Lee, Y. M., Madigan, R., Merat, N., & Romano, R. (2022). Explaining unsafe pedestrian road crossing behaviours using a psychophysics-based gap acceptance model. *Saf. Sci.,, 154*(105837), Article 105837.

Tom, A. and Granié, M.-A. (2011). Gender differences in pedestrian rule compliance and visual search at signalized and unsignalized crossroads. *Accident Analysis & Prevention*, 43(5):1794–1801.

Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R., Stansifer, R., and Castellan, N. J. (1979). Tri-level study of the causes of traffic accidents: final report. executive summary. Technical report, Indiana University, Bloomington, Institute for Research in Public Safety.

Venturino, M. and Wells, M. J. (1990). Head movements as a function of field-of-view size on a helmet-mounted display. *Proceedings of the Human Factors Society Annual Meeting*, 34(19):1572–1576.

Wells, M. J. and Venturino, M. (1990). Performance and head movements using a helmet-mounted display with different sized fields-of-view. *Optical Engineering*, 29 (8):870–877.

Wessels, M., Zähme, C., & Oberfeld, D. (2023). Auditory information improves time-to- collision estimation for accelerating vehicles. *Current Psychology, 42*(27), 23195–23205.

Whitebread, D. and Neilson, K. (2001). The contribution of visual search strategies to the development of pedestrian skills by 4-11 year-old children. *The British journal of educational psychology*, 70 Pt 4:539–57.

WHO, W. H. O. (2023). Global status report on road safety 2023. URL: https://www.who.int/ publications/i/item/9789240086517. Accessed on 25 November 2024.

Yang, Y., Lee, Y. M., Madigan, R., Solernou, A., & Merat, N. (2024). Interpreting pedestrians' head movements when encountering automated vehicles at a virtual crossroad. *Transportation research part F: traffic psychology and behaviour, 103*, 340–352.