How do drivers overtake pedestrians? Evidence from field test and naturalistic driving data

Accident Analysis and Prevention 139 (2020) 105494

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

How do drivers overtake pedestrians? Evidence from field test and naturalistic driving data

Alexander Rasch*, Gabriele Panero, Christian-Nils Boda, Marco Dozza

Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Hörselgången 4, 41756, Göteborg, Sweden

ARTICLE INFO

Keywords:
Pedestrian safety
Comfort zone
Overtaking
Euro NCAP
Driver behavior
Bayesian regression modeling

ABSTRACT

For pedestrians, the risk of dying in a traffic accident is highest on rural roads, which are often characterized by a lack of sidewalks and high traffic speed. In fact, hitting the pedestrian during an overtaking attempt is a common crash scenario. To develop active safety systems that avoid such crashes, it is necessary to understand and model driver behavior during the overtaking maneuvers, so that system interventions are acceptable because they happen outside drivers' comfort zone. Previous modeling of driver behavior in interactions with pedestrians primarily focused on road crossing scenarios. The aim of this study was, instead, to address pedestrian-overtaking maneuvers on rural roads. We focused our analysis on how drivers adjust their behavior with respect to three safety metrics (in order of importance): 1) minimum lateral clearance when passing the pedestrian, 2) overtaking speed at that moment, and 3) the time-to-collision at the moment of steering away to start the overtaking maneuver.

The influence of three factors on the safety metrics was investigated: 1) walking direction (same as the overtaking vehicle or opposite), 2) walking position (on the edge of the vehicle lane or 0.5 m away from the edge on the paved shoulder), and 3) oncoming traffic (absent or present). Seventy-seven overtaking maneuvers in France from the naturalistic driving study UDRIVE and 297 maneuvers in Sweden from field tests were analyzed. Bayesian regression was used to model how minimum lateral clearance and overtaking speed depended on the three factors. Results showed that drivers maintained smaller minimum lateral clearance and lower overtaking speed when the pedestrian was walking in the opposite direction, on the lane edge, or when oncoming traffic was present. Minimum lateral clearance and time-to-collision were only weakly correlated with overtaking speed. The regression models predicted distributions similar to those actually observed in the data. The time-to-collision at the moment of steering away was comparable in value to the time-to-collision used by Euro NCAP for testing active safety systems in car-to-pedestrian longitudinal scenarios since 2018.

This study is the first to analyze driver behavior when overtaking pedestrians, based on field test and naturalistic driving data. Results suggest that pedestrian safety is particularly endangered in situations when the pedestrian is walking opposite to traffic, close to the lane, and when oncoming traffic is present. The Bayesian regression models from this study can be used in active safety systems to model drivers' comfort in overtaking maneuvers.

1. Introduction

Globally, pedestrians represent 22 % of all traffic fatalities, albeit with large regional differences (World Health Organization, 2018). In the European Union (EU), pedestrians account for 21 % of all traffic fatalities. The fatality risk to pedestrians is highest on rural roads, which often lack sidewalks and have high traffic speed (European Road Safety Observatory, 2016; Zegeer and Bushell, 2012).

A botched overtaking of a pedestrian walking along the road is a common crash scenario. Although it is defined differently in different national crash statistics, the following figures should provide some idea of its prevalence. In the EU, this scenario accounts for 10 % of all pedestrian fatalities (Lübbe, 2015). Differences inside the EU do exist: in Great Britain this rate is 24 %, in Germany 5%, and in France 8% (Wisch et al., 2013). In Sweden, the STRADA crash database (Howard and Linder, 2014) reports that between 2003 and 2010, 8% of all pedestrian injuries in crashes with motorized traffic occurred when the pedestrian was walking on the right side of the road. In the US, this

* Corresponding author.
E-mail addresses: alexander.rasch@chalmers.se (A. Rasch), gabriele.panero@veoneer.com (G. Panero), christian-nils.boda@chalmers.se (C.-N. Boda), marco.dozza@chalmers.se (M. Dozza).

https://doi.org/10.1016/j.aap.2020.105494
Received 11 April 2019; Received in revised form 5 March 2020; Accepted 5 March 2020
Available online 20 March 2020

same scenario is the second most predominant pedestrian crash scenario, accounting for 27 % of all vehicle-pedestrian crashes (Yanagisawa et al., 2017). In China, the scenario of a pedestrian moving along the road (in the center or beside a two-lane road) accounted for 26 % of the pedestrian fatalities between 2011 and 2014 (Chen et al., 2015).

Pedestrian safety in overtaking scenarios on rural roads can be addressed by improving: 1) infrastructure, 2) traffic regulation, and 3) in-vehicle safety systems. Infrastructure solutions typically consist of separated walking zones such as sidewalks (Laird et al., 2013; World Health Organization, 2018). On the other hand, traffic regulations focus on the lateral clearance during overtaking maneuvers and give recommendations about pedestrian's walking behavior. In fact, the Vienna convention on road traffic (signed and ratified by 78 countries) recommends that single pedestrians on a shared road be advised to walk in the opposite direction of the traffic (United Nations, 1968). Swedish and French traffic regulations request that pedestrians walk on the left side, i.e. facing the traffic, when there is no sidewalk (Trafikförordning, Ch. 7, § 1, and Code de la route, Art. R.412-36, respectively).

The development of in-vehicle safety systems is influenced by the European new car assessment program (Euro NCAP), which has been testing active safety systems for pedestrian protection since 2016 (Schram et al., 2015; Van Ratingen et al., 2016). Since 2018, the Euro NCAP test protocol includes the longitudinal car-to-pedestrian adult (CPLA) scenario, relating to our overtaking scenario (Euro NCAP, 2019). The program's testing includes intervention systems such as autonomous emergency braking (AEB) and warning systems such as forward collision warning (FCW). Automatic emergency steering (AES) systems are planned to be included for assessment in 2020 (Euro NCAP, 2018). Both intervention and warning systems need to be precisely tuned, so that drivers do not receive alerts or interventions that are either too early or unnecessary (Brännstrom et al., 2013; Lubbe and Davidsson, 2015; Lubbe and Rosén, 2014). Recent studies of automated steering maneuvers involving pedestrian interactions suggest that steering away earlier and maintaining longer lateral distances (compared to manual driving) results in better driver acceptance (Abe et al., 2018).

Studies have suggested that modeling driver behavior—and in particular, driver's comfort— would provide information that could be used to make interventions and warnings more acceptable to drivers (Brännstrom et al., 2013; Summala, 2007). In 2011, Ljung Aust and Engström presented a framework for modeling drivers' comfort zone, which represents the spatiotemporal region in which the driver does not feel discomfort (Ljung Aust and Engström, 2011). The underlying assumption is that active safety warnings and interventions will be intrinsically more acceptable when drivers have surpassed their comfort zone boundaries (Brännström et al., 2014).

To assess drivers' comfort zone when they are overtaking a cyclist, previous studies (Dozza et al., 2016; Kovaceva et al., 2018), proposed that overtaking maneuvers be divided into four phases: approaching, steering away, passing, and returning. Because no similar work has been done for pedestrian-overtaking maneuvers, in this study we used and further refined this definition. To date, existing research on driver-pedestrian interaction has primarily focused on intersection scenarios (Lubbe and Davidsson, 2015; Lubbe and Rosén, 2014; Ren et al., 2016), while overtaking scenarios have not been addressed yet (to our knowledge). This work addresses that research gap.

The main aim of this study was to leverage field test (FT) and naturalistic driving (ND) data to investigate how three different factors influenced minimum lateral clearance and vehicle speed during pedestrian-overtaking maneuvers. We hypothesized that these measures would be smaller when: 1) the pedestrian was traveling in the same direction as the vehicle (instead of the opposite direction), 2) the pedestrian walked on the lane edge (instead of 0.5 m away from the edge on the paved shoulder), 3) there was an oncoming vehicle meeting the overtaking driver (instead of no vehicle). A secondary aim of this study was to fit similar Bayesian regression models for each of the two datasets with respect to the three factors in order to predict how drivers choose minimum lateral clearance and speed when overtaking a pedestrian. We further g vehicle meeting the overtaking driver (instead of no vehicle). A secondary aim of this study was to fit similar Bayesian regression models for each of the two datasets with respect to the three factors in order to predict how drivers choose minimum lateral clearance and speed when overtaking a pedestrian. We further discuss possible advantages of Bayesian models for modeling driver behavior from datasets collected in different environments and for improving active safety systems. As a tertiary aim, we also measured TTC when drivers steered away, since this could be compared to the TTC used by Euro NCAP.

2. Material and methods

This study analyzed how drivers overtake pedestrians using two datasets: 1) ND data from the UDRIVE study (Barnard et al., 2016), and 2) FT data from an ad hoc experiment (Fig. 1). Since the two datasets were obtained in two fundamentally different set ups (ND unobtrusively, FT in a planned experiment), and different environments, we will analyze them separately in the results section. ND data have higher ecological validity but are confounded by a large variety of environmental factors (Bärgman, 2016). FT data are controlled for environmental factors at the expense of their ecological validity (Boda, 2017). By comparing results from both datasets (in the discussion section), we build cumulative evidence to support our results.

Overtaking maneuvers were initially separated into three driver strategies: *accelerative* (reducing the speed to that of the pedestrian and waiting until the oncoming vehicle has gone by, then accelerating past the pedestrian), *flying* (maintaining speed and passing the pedestrian ahead of the oncoming vehicle) and *piggybacking* (imitating a lead vehicle's overtaking behavior). These definitions were inspired by previous research on overtaking of vehicles and cyclists (Matson and Forbes, 1938). No accelerative maneuvers were observed in our datasets due to the great difference in speed between vehicle and pedestrian. Further, piggybacking maneuvers were excluded from the analysis, as well as maneuvers from heavy vehicles. We excluded

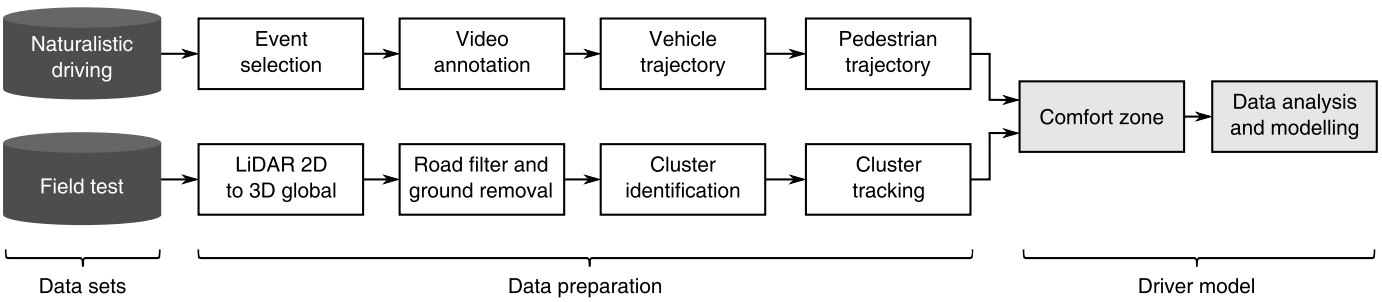

Fig. 1. Data processing steps to prepare and utilize naturalistic driving and field test data for driver modeling.

piggybacking maneuvers because a second vehicle was involved which could have influenced the driver's behavior. Similarly, heavy vehicles were ignored because Dozza et al. (2016) showed that truck drivers behave differently than car drivers when overtaking cyclists and we were concerned that the behavior difference may apply to pedestrian-overtaking maneuvers as well. Thus, only flying overtaking maneuvers from personal vehicles are part of this analysis, although some of these maneuvers included mild braking.

2.1. The UDRIVE naturalistic driving study

This study obtained ND data collected in South Eastern France from 30 passenger cars from UDRIVE, the largest ND study in Europe. The drivers were known and had given consent for their data to be used in this study. Requirements on participation in the study were based on experience (minimum annual driving distance of 10,000 km), age (three age groups: 18–25, 26–45 and 46–70) and gender (minimum of 40 % per gender). The vehicles used in the UDRIVE study were Renault Megane, Renault Clio 3 and Renault Clio 4. Each vehicle was equipped with a data acquisition system which recorded data from 7 cameras, IMU sensors, GPS, MobilEye smart camera, CAN data and sound level. For further information about the collection of the data, please consult the overview of the main results of the UDRIVE project by Van Nes et al. (2017). (Barnard et al., 2016; Lai et al., 2013)

The signals from the UDRIVE dataset used in this study included GPS location, CAN data, video data (indoor and forward-facing cameras), and MobilEye data. The MobilEye camera pointed to the forward roadway, providing relative pedestrian position and speed within its field of view (± 25°). Overtaking events were selected and annotated, adapting a procedure from Kovaceva et al. (2018).

2.1.1. Event selection

We selected segments, i.e. time series of data, which fulfilled the following criteria: 1) MobilEye detected a single pedestrian, 2) the vehicle speed was above 20 km/h, and 3) the lateral position of the pedestrian in the vehicle reference frame (i.e. from the MobilEye camera perspective) when the lateral acceleration of the vehicle was at its maximum was greater than the pedestrian's average lateral position.

2.1.2. Video annotation

The video of each selected segment was manually analyzed to verify that a pedestrian-overtaking maneuver had taken place. The reasons for excluding segments were: MobilEye obstacle misdetection (typically, traffic signs classified as pedestrians), pedestrian walking on a dedicated path (such as a sidewalk), or presence of another road user in the scene (usually, vehicles parked close to the pedestrian).

Each segment containing an overtaking maneuver was further annotated by two analysts during a time span of about two weeks. Video annotation from the driver facing in-cabin camera in combination with the steering angle signal from CAN, identified the start of the overtaking maneuver (when the driver started to rotate the steering wheel to evade the collision with the pedestrian). Video annotation from the front view camera classified 1) pedestrian walking direction (same or opposite to the traffic), 2) pedestrian walking position (lane edge or road shoulder), 3) oncoming traffic (presence or absence), and 4) the pedestrian velocity (categorized as standing, 0 km/h; walking, 5 km/h; or running, 9 km/h). To estimate comfort zone, we calculated the vehicle and pedestrian trajectories (Fig. 2).

2.1.3. Vehicle trajectory

Because of its limited field of view, the MobilEye camera eventually lost track of the pedestrian during the overtaking maneuver. Fig. 2 shows the kinematic variables used to reconstruct the vehicle's and pedestrian's trajectories after the pedestrian disappeared from the camera's field of view.

The vehicle trajectory in two-dimensional space was computed by integrating the vehicle speed, $\dot{\mathbf{x}}_{\text{veh}}$, expressed as follows:

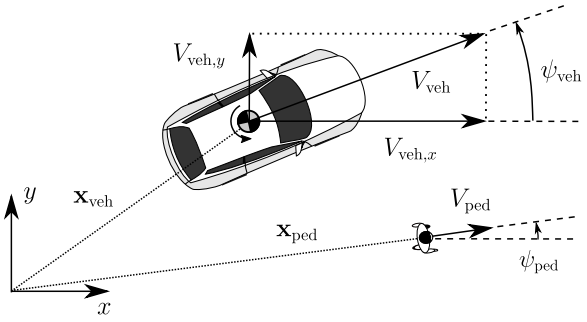

Fig. 2. Vehicle and pedestrian kinematic variables in the global frame.

$$\dot{\mathbf{x}}_{\text{veh}} = \begin{bmatrix} V_{\text{veh},x} \\ V_{\text{veh},y} \end{bmatrix} = V_{\text{veh}} \begin{bmatrix} \cos(\psi_{\text{veh}} + \alpha(\delta)) \\ \sin(\psi_{\text{veh}} + \alpha(\delta)) \end{bmatrix}, \tag{1}$$

where $V_{\text{veh},x}$ and $V_{\text{veh},y}$ are the longitudinal (x) and lateral speeds (y) of the vehicle, respectively. V_{veh} is the magnitude of the vehicle speed vector, α the sideslip angle, δ the wheel steer angle, and ψ_{veh} the heading angle (Rajamani, 2012). Since none of the events contained excessive steering input from the driver, α was neglected. V_{veh} and $\dot{\psi}_{\text{veh}}$ were obtained from the CAN bus and the onboard gyroscope, respectively. To estimate the position of the vehicle, first the heading angle rate $\dot{\psi}_{\text{veh}}$ was integrated; then the longitudinal and lateral values in the global frame could be estimated with one more trapezoidal integration.

2.1.4. Pedestrian trajectory

MobilEye provided the pedestrian position in the vehicle reference frame while the pedestrian was within the field of view. Once the pedestrian was no longer visible, we calculated the pedestrian trajectory from the traveled distance and converted the trajectory to the global frame.

The traveled distance p_{ped} was obtained by forward Euler integration:

$$p_{\text{ped}}[t_k] = p_{\text{ped}}[t_{k-1}] + V_{\text{ped}}\,(t_k - t_{k-1}). \tag{2}$$

Eq. (2) includes the estimated traveled distance at time t_k, $p_{\text{ped}}[t_k]$, based on the previous distance at time t_{k-1} (assuming the pedestrian velocity V_{ped} to be constant during the whole maneuver). At time t_0, when the pedestrian left the field of view, the traveled distance ($p_{\text{ped}}[t_0]$) was set to zero. Because MobilEye did not always measure of pedestrian velocities accurately, the velocity it provided was compared with the video-annotated velocity. For 58 out of the 83 events, the annotated velocity differed by 2 km/h or more from the MobilEye velocity and was taken instead. Due to sudden deviations in the MobilEye-detected pedestrian positions, the pedestrian's trajectory during the overtaking maneuver was approximated with a line. The slope of the line, corresponding to the heading angle of the pedestrian ψ_{ped} in the global frame, was estimated with the random sample consensus (RANSAC) algorithm (Fischler and Bolles, 1981). The line was designed to include the mean position in the x and y directions, $\bar{\mathbf{x}}_{\text{ped}}$, obtained from the MobilEye detected positions. Eq. (3) shows how the pedestrian position in the global frame was calculated from the traveled distance and the RANSAC-estimated heading angle.

$$\mathbf{x}_{\text{ped}} = \bar{\mathbf{x}}_{\text{ped}} + p_{\text{ped}} \begin{bmatrix} \cos(\psi_{\text{ped}}) \\ \sin(\psi_{\text{ped}}) \end{bmatrix} \tag{3}$$

2.2. Field test data from ad hoc experiment

Field test data were collected in an experiment on a public road in Sweden. We therefore equipped a pedestrian with a data logger to measure the overtaking maneuvers by drivers.

Table 1
Sensors connected to the data logger.

Sensor	Product	Sample rate	Resolution
LiDAR	Hokuyo UXM-30LAH-EWA	20 Hz	Range: 1 mm, angular: 0.125°
IMU	PhidgetSpatial 1044_0	250 Hz	Accelerometer: 76.3 μg, Gyroscope: 0.02°/s, Magnetometer: 3 mG
Camera	Creative Live! Cam Sync HD	15 fps	640 × 480 px
GPS receiver	Globalsat BU-353S4	1 Hz	Accuracy < 2.5 m 2D root mean square
Flag button	ITW 59–111 push button	10 Hz	N/A

2.2.1. Hardware

Table 1 lists the hardware components of the data logger used in the field experiment. The main computing unit was a Raspberry Pi 3 Model B, running Ubuntu 16.04.4 LTS. The Raspberry Pi was chosen because of its low cost relative to its computation power and its previous usage in research (Ambrož, 2017; Dozza et al., 2017). The Adafruit DS1307 Real Time Clock was used to keep the correct time after a reboot. A USB drive stored the data. The pedestrian was equipped with a flag button (to add time marks to the recorded data—typically when being overtaken by a vehicle) and a GPS receiver (to estimate global position). A web camera was added to record video. A LiDAR sensor with a 190-degree field of view provided a two-dimensional scan of the environment around the pedestrian. An inertial measurement unit (IMU) was added to estimate orientation. The camera, LiDAR, and IMU were cased together, so that all these sensors experienced the same kinematics. Together, these five sensors comprised the data logging system worn by the pedestrian. A detailed description of the hardware and the 3D models for the casing parts is publicly available[1] .

2.2.2. Software

The data logger software was based on a robot operating system (ROS) package and written in Python. ROS is an open-source, modular framework for software development in robotics applications or any other hardware platforms which involve reading sensor values or actuation (Quigley et al., 2009). Fig. 3 shows the ROS package architecture of the data logger. The software was based on a central node, *datalogger*, implementing a state machine with the states *idle*, *record*, and *shutdown*. The data logger was controlled by an HTML interface which could be accessed wirelessly (e.g., from a phone).

2.2.3. Experimental protocol

Field data were collected on two days in March 2018 in the early afternoon, with clear sight conditions and temperatures around 2 ^{0}C, on the Tuvevägen road in Tuve, Sweden (GPS coordinates of the center of the stretch of road used are 57.763763° longitude, 11.936956° latitude). The road was about 1 km long, straight, and had two lanes (each 3.5 m wide). The speed limit was 70 km/h. The pedestrian wore the logging system attached at the hip (Fig. 4a) and walked according to four scenarios (Fig. 4b and c) created by combining two factors: 1) whether or not the pedestrian walked facing traffic and 2) whether the pedestrian walked on the line delimiting the lane or on the road shoulder (0.5 m away from the line).

In all scenarios, the LiDAR was facing the road. Hence, when switching pedestrian direction, the data logger needed to be transferred to the opposite hip. The camera was facing the overtaking vehicles in their approaching phase. For safety, a team was close by and connected to the pedestrian via phone at all times during data collection.

2.2.4. Data processing

To compensate for the rotation movement of the LiDAR and to sanitize the data, we followed four main steps, according to the workflow represented in Fig. 1.

2.2.4.1. LiDAR 2D to 3D global. A Madgwick filter for an attitude and heading reference system (Madgwick, 2010) was used to estimate the LiDAR orientation. The filter takes accelerometer, gyroscope and magnetometer readings (from the IMU) as input and calculates the orientation in a quaternion representation, using an optimized gradient descent algorithm. The LiDAR and IMU data were then processed with the Point Cloud Library (Rusu and Cousins, 2011) to obtain a 3D representation of the data. Artefacts from the yaw movement induced

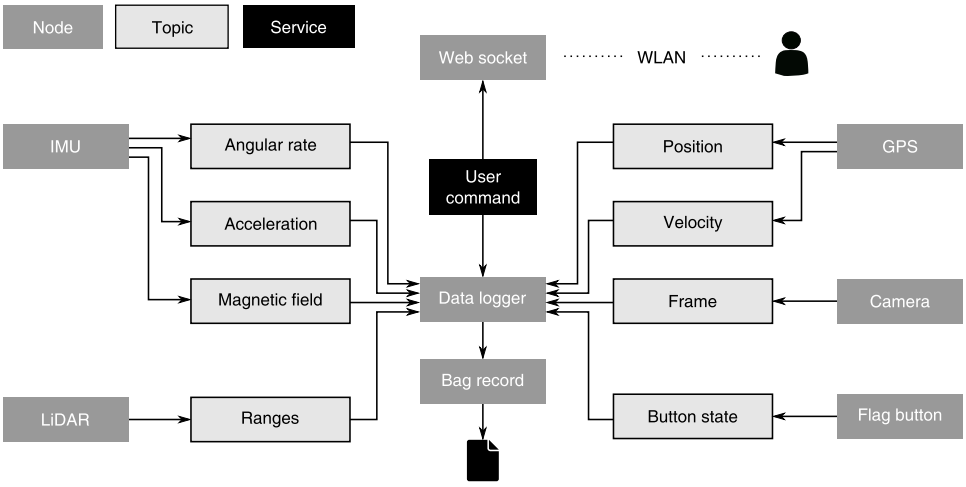

Fig. 3. Data logger software overview; The ROS architecture allows a user to send commands via a web site and record all topics in a ROS bag file.

[1] https://github.com/ruvigroup/div_datalogger.

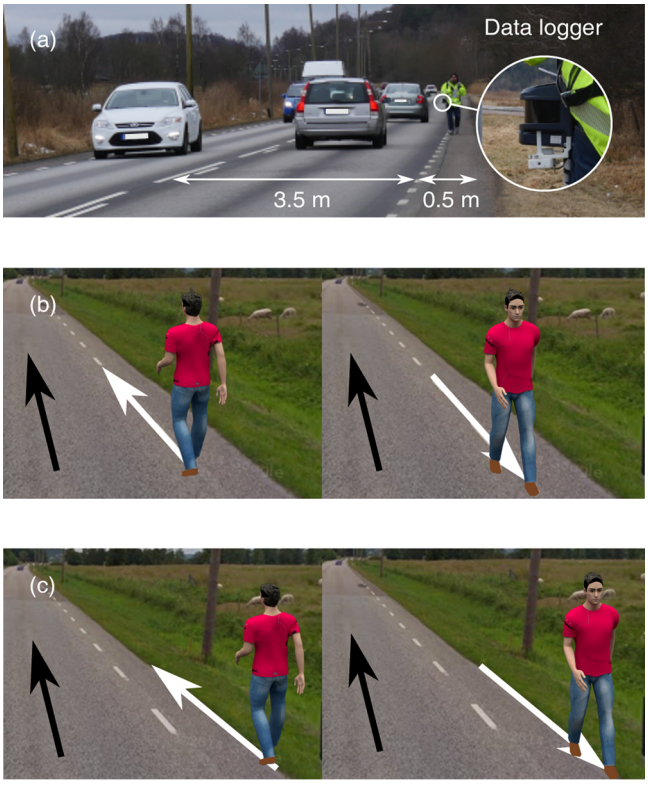

Fig. 4. Field test data collection. Panel (a) shows the field test scenario. Panels (b) and (c) clarify the pedestrian's walking position: shoulder and line, respectively. Traffic and walking directions are indicated with black and white arrows, respectively.

by walking were cancelled out.

2.2.4.2. Road filter and ground removal. The point cloud contained a substantial amount of noise, due to the environment (mainly vegetation next to the road and reflection from the road itself). The noise due to vegetation was removed by taking away the points in the rectified and transformed point cloud outside of eight meters, measured perpendicular to the walking direction of the pedestrian. The point cloud was also limited in the vertical direction to filter out any reflection from the ground.

2.2.4.3. Cluster identification. The resulting point cloud contained detections of multiple vehicles—for example, when there was a vehicle piggybacking or oncoming. To identify individual vehicles, we used a k-d tree object as the search method for a Euclidean cluster extraction algorithm (Rusu, 2010).

2.2.4.4. Cluster tracking. The clusters identified on individual frames were then tracked across time and assigned unique identifiers. Using RANSAC, we estimated the cluster speeds, which we combined with the GPS speed of the pedestrian to retrieve the absolute speed of each vehicle.

Since the speed of overtaking vehicles was assumed to be constant during the maneuver, vehicles were automatically classified as piggybackers when the time headway to the previous vehicle was lower than 3 s. Each overtaking event was checked to determine if an oncoming vehicle was present in the scene (from 20 m behind the pedestrian to 120 m in front). Cluster size was used to filter out large vehicles like trucks and buses.

2.3. Comfort zone

Comfort zone was expressed in terms of minimum lateral clearance (MLC), i.e. the minimum distance between the vehicle and pedestrian, shown in Fig. 5. The figure also shows the four phases of the overtaking maneuver: 1) *approaching*, 2) *steering away*, 3) *passing*, and 4) *returning*, adapted from Dozza et al.'s 2016 work. For the ND data, time-to-collision (TTC) to the pedestrian at the beginning of the steering away phase was computed.

2.4. Data analysis and modeling

Using Bayesian regression models (BRMs), we analyzed MLC and overtaking speed in order to address our hypotheses and to derive models which could be used in active safety systems. BRMs approximate the posterior density of a model within certain parameters conditioned on data by making use of MCMC (Markov chain Monte-Carlo) sampling methods, and are therefore able to express the uncertainty of parameter estimates (Feinberg and Gonzalez, 2012). BRMs can be fitted to data with arbitrary distributions and a nested structure or hierarchy of populations (Bürkner, 2017; Hoff et al., 2006; Kruschke, 2014; Morando, 2019). In the discussion section, we will argue how using a Bayesian modeling approach could improve driver adaptation in active safety systems. TTC was used to relate our results to the CPLA scenarios used in Euro NCAP for the evaluation of active safety systems.

The R Project's software was used to fit the BRMs, using the package *brms*, version 2.8.0 (Bürkner, 2017). The three binary factors considered were: pedestrian walking direction X_{dir} (0 = opposite, 1 = same), pedestrian walking position X_{pos} (0 = shoulder, 1 = line) and oncoming traffic X_{onc} (0 = absent, 1 = present).

At first, a model including all possible two- and three-way interactions between the three factors was fitted. The model was then improved by excluding the parameters with only a small impact on the response. Leave-one-out cross-validation (LOOCV)—and in particular the expected predictive accuracy—was used to verify the improvement (Vehtari et al., 2017). Each BRM was fitted in *brms* using the No-U-Turn sampler (Hoffman and Gelman, 2011) with eight chains and weakly informative (default) prior distributions for the parameters (Bürkner, 2017). Each chain comprised 10,000 samples; the first 5000 were used for warm-up by the sampler and thereafter discarded from the analysis. The number of iterations was chosen large enough to guarantee convergence of the chains. We verified the convergence by inspection of

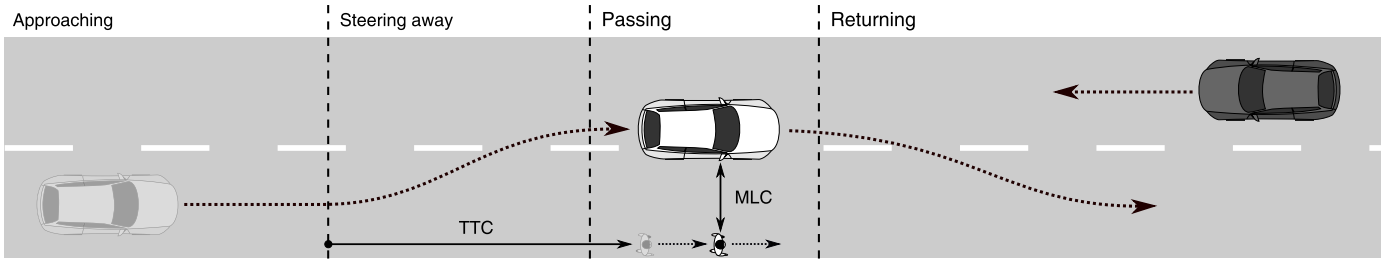

Fig. 5. Comfort zone measures minimum lateral clearance (MLC) and time to collision (TTC) during an overtaking maneuver (here shown with oncoming traffic). The overtaking maneuver is divided into four phases.

the trace plots of the chains, and an *Rhat* value close to 1 (Bürkner, 2017). Drawing samples from the fitted models allowed us to quantify the differences between the two outcomes for each binary factor in the form of 95 % highest density intervals (HDIs).

The 95 % HDIs of the model responses obtained from the Bayesian models describe the uncertainty in MLC and overtaking speed in terms of credibility. The 95 % HDI can be compared to a region of practical equivalence (ROPE), an interval that specifies "practical" null values, and on that basis the null hypothesis can be either accepted, rejected, or left unresolved, in an alternative to null hypothesis significance testing from frequentist statistics. According to the methods for Bayesian hypothesis testing proposed by Kruschke (2018), only if the HDI falls completely outside of the ROPE can the null hypothesis be rejected, and only if the HDI falls completely inside the ROPE can the null hypothesis be accepted.

2.4.1. Bayesian model of minimum lateral clearance

Since drivers always maintained a positive MLC from the pedestrian, resulting in a skewed data distribution, a log-normal distribution was chosen to model MLC. In Eq. (4), the model for the i th response variable $\mathrm{MLC}_{j,i}$ of the dataset with index j, with $j \in \{\mathrm{FT}, \mathrm{ND}\}$, is shown:

$$\mathrm{MLC}_{j,i} \sim \mathrm{Lognormal}(\mu_{\mathrm{MLC},j,i}, \sigma_{\mathrm{MLC},j}), \tag{4}$$

where $\mu_{\mathrm{MLC},j,i}$ is the log-normal mean of the model of the i th response variable for dataset j, called the predictor. $\sigma_{\mathrm{MLC},j}$ is the standard deviation of the log-normal distribution, assumed constant over all responses of dataset j. The predictor for the FT dataset is:

$$\mu_{\mathrm{MLC,FT},i} = \boldsymbol{X}_{\mathrm{MLC,FT},i}\, \boldsymbol{\beta}_{\mathrm{MLC}}, \tag{5}$$

where $\boldsymbol{\beta}_{\mathrm{MLC}}$ is the vector of population-level parameters and $\boldsymbol{X}_{\mathrm{MLC,FT}}$ the corresponding design matrix. $\boldsymbol{\beta}_{\mathrm{MLC}}$ contains the parameters to be fitted, and (for the full model with all interactions) is defined as:

$$\boldsymbol{\beta}_{\mathrm{MLC}} = [\beta_0\ \beta_{\mathrm{dir}}\ \beta_{\mathrm{pos}}\ \beta_{\mathrm{onc}}\ \beta_{\mathrm{dir*pos}}\ \beta_{\mathrm{dir*onc}}\ \beta_{\mathrm{pos*onc}}\ \beta_{\mathrm{dir*pos*onc}}]^T, \tag{6}$$

Where the subscript 0 stands for the intercept of the model, *dir* for walking direction, *pos* for walking position and *onc* for oncoming traffic. The * operator marks an interaction between factors. Each row i of the design matrix, $\boldsymbol{X}_{\mathrm{MLC,FT},i}$, is accordingly defined as:

$$\boldsymbol{X}_{\mathrm{MLC,FT},i} = [1\ X_{\mathrm{dir},i}\ \ X_{\mathrm{pos},i}\ \ X_{\mathrm{onc},i}\ \ X_{\mathrm{dir},i}\ X_{\mathrm{pos},i}\ \ X_{\mathrm{dir},i}\ X_{\mathrm{onc},i}\ \ X_{\mathrm{pos},i}\ \ X_{\mathrm{onc},i}\ \ X_{\mathrm{dir},i}\ X_{\mathrm{pos},i}\ \ X_{\mathrm{onc},i}]. \tag{7}$$

The design matrix contains the factor values for each row of data, i, where X_{dir} represents the factor walking direction (0 = opposite, 1 = same), X_{pos} the factor walking position (0 = shoulder, 1 = line), and X_{onc} the factor oncoming traffic (0 = absent, 1 = present).

Since the ND dataset—in contrast to the FT dataset—contained multiple events from the same drivers, driver identity (ID) was included as a group-level effect (Bürkner, 2018), resulting in the predictor:

$$\mu_{\mathrm{MLC,ND},i} = \boldsymbol{X}_{\mathrm{MLC,ND},i}\, \boldsymbol{\beta}_{\mathrm{MLC}} + \boldsymbol{Z}_{\mathrm{MLC,ID},i}\boldsymbol{u}_{\mathrm{MLC,ID}},\quad \boldsymbol{u}_{\mathrm{MLC,ID}} \sim N(0,\ \sigma^2_{\mathrm{MLC,ID}}\boldsymbol{I}). \tag{8}$$

In Eq. (8), $\boldsymbol{X}_{\mathrm{MLC,ND},i}$ is the i th row of the design matrix and defined in accordance to Eq. (7). $\boldsymbol{u}_{\mathrm{MLC,ID}}$ contains the group-level effect of each driver (with driver identity *ID*) on the intercept of the predictor. $\boldsymbol{u}_{\mathrm{MLC,ID}}$ is modelled to be sampled from a zero-centered, multivariate, normal distribution with standard deviation $\sigma_{\mathrm{MLC,ID}}$, and $\boldsymbol{Z}_{\mathrm{MLC,ID},i}$ is the i th row of the corresponding design matrix $\boldsymbol{Z}_{\mathrm{MLC,ID}}$.

2.4.2. Bayesian model of overtaking speed

The speed of the vehicle V_{veh} when overtaking the pedestrian was modeled from ND and FT data. We chose a Student's t-distribution to include more outliers, under the assumption that the data distribution was approximately symmetrical and had a fatter tail than a normal distribution.

$$V_{\mathrm{veh},j,i} \sim \mathrm{Student{-}t}(\mu_{V,j,i},\ \sigma_{V,j},\ \nu_{V,j}), \tag{9}$$

where $\mu_{V,j,i}$ is the location, $\sigma_{V,j}$ the scale, and $\nu_{V,j}$ the degrees of freedom.

The predictors of the BRMs for FT and ND data are described by Eqs. (10) and (11), respectively.

$$\mu_{V,\mathrm{FT},i} = \boldsymbol{X}_{V,\mathrm{FT},i}\, \boldsymbol{\beta}_V, \tag{10}$$

$$\mu_{V,\mathrm{ND},i} = \boldsymbol{X}_{V,\mathrm{ND},i}\, \boldsymbol{\beta}_V + \boldsymbol{Z}_{V,\mathrm{ID},i}\boldsymbol{u}_{V,\mathrm{ID}},\quad \boldsymbol{u}_{V,\mathrm{ID}} \sim N(0,\ \sigma^2_{V,\mathrm{ID}}\boldsymbol{I}). \tag{11}$$

Both predictors include the i th row of the design matrix $\boldsymbol{X}_{V,j}$ and parameters $\boldsymbol{\beta}_V$, analogous to the MLC model. Eq. (11) also includes the group-level effect of the driver identity, $\boldsymbol{u}_{V,\mathrm{ID}}$, and the corresponding design matrix $\boldsymbol{Z}_{V,\mathrm{ID}}$.

3. Results

In the following, we will present an overview of both datasets and the results for MLC and overtaking speeds across conditions and for each data set.

3.1. Data overview

ND data from UDRIVE provided 83 pedestrian-overtaking maneuvers. FT data provided 481 pedestrian-overtaking maneuvers. Six maneuvers from ND data and 184 from FT data were excluded from the analysis because they included either vehicles piggybacking or heavy vehicles. The ND data used for analysis consisted of 77 maneuvers, performed by 23 drivers. The FT data used for analysis consisted of 297 maneuvers, which were assumed to be performed by individual drivers. MLC and overtaking speed are represented for ND and FT data in Tables 2 and 3, respectively.

Table 2
Means (with standard deviations within parentheses) for minimum lateral clearance (MLC) and overtaking speed from naturalistic driving data.

Walking direction	opposite				same			
Walking position	shoulder		line		shoulder		line	
Oncoming traffic	absent	present	absent	present	absent	present	absent	present
Number of events	3	8	14	5	3	9	20	15
MLC [m], mean (std)	1.46 (0.45)	1.40 (0.43)	0.98 (0.35)	0.89 (0.23)	1.82 (0.36)	1.60 (0.67)	1.07 (0.47)	0.90 (0.37)
Overtaking speed [km/h], mean (std)	63.3 (10.9)	55.4 (8.7)	59.4 (19.4)	47.3 (13.6)	60.0 (18.9)	60.4 (10.1)	49.1 (11.7)	45.9 (12.9)

Table 3
Means (with standard deviations within parentheses) for minimum lateral clearance (MLC) and overtaking speed from field test data.

Walking direction	opposite				same			
Walking position	shoulder		line		shoulder		line	
Oncoming traffic	absent	present	absent	present	absent	present	absent	present
Number of events	31	17	55	71	23	18	54	28
MLC [m], mean (std)	1.87 (0.44)	1.57 (0.29)	1.54 (0.44)	1.28 (0.34)	2.00 (0.47)	1.83 (0.47)	1.76 (0.44)	1.52 (0.33)
Overtaking speed [km/h], mean (std)	63.6 (7.2)	58.2 (8.1)	59.2 (9.2)	56.0 (7.3)	60.1 (7.5)	57.2 (7.2)	59.6 (9.2)	57.9 (10.0)

3.2. Minimum lateral clearance

Fig. 6 shows the influence of the pedestrian walking direction, pedestrian walking position, and the presence of oncoming traffic on MLC. The Pearson correlation coefficient revealed a small positive correlation between MLC and overtaking speed for FT data (r = 0.25). For ND data, the correlation was similar (r = 0.22).

All three binary factors influenced MLC within a 95 % HDI in FT data. For ND data, MLC trends (in relation to the three factors) were consistent with FT data, albeit not within a 95 % HDI. Table 4 shows the estimates of the optimized model (after the exclusion of all interactions). The model parameters for the full MLC model with all interactions for ND and FT data are included in Table A1 in Appendix A.

LOOCV verified that the optimized MLC models in Table 4 were indeed better fits than the complete models, as indicated by the difference in expected predictive accuracy. The accuracy was higher by 2.9 (with 1.2 standard error) and 3.2 (with 1.1 standard error) for FT and ND data, respectively, compared to the full models.

Table 5 shows the effects on MLC of the differences in possible outcomes for the three factors, as medians—for ND and FT data and their models. The table also reports the 95 % HDIs of the models. The validity of our models is supported by the fact that all median differences of all effects for ND and FT data fall within the 95 % HDI of our models.

Table 4
Parameter distributions of the optimized models for minimum lateral clearance, fitted from naturalistic driving (ND) and field test (FT) data. Medians are reported together with the lower and upper limits of the 95 % highest density interval (HDI), l-95 % HDI and u-95 % HDI, respectively. $\sigma_{MLC,ID}$ is the standard deviation of the group-level effect driver identity (ID), and σ_{MLC} the standard deviation parameter of the log-normal distribution.

Parameter	ND model			FT model		
	Median	l-95 % HDI	u-95 % HDI	Median	l-95 % HDI	u-95 % HDI
β_0	0.18	−0.03	0.38	0.48	0.43	0.53
β_{dir}	0.04	−0.15	0.22	0.14	0.08	0.21
β_{pos}	−0.37	−0.53	−0.22	−0.14	−0.18	−0.09
β_{onc}	−0.11	−0.31	0.09	−0.16	−0.22	−0.09
$\sigma_{MLC,ID}$	0.17	0.00	0.36	N/A	N/A	N/A
σ_{MLC}	0.38	0.31	0.46	0.27	0.25	0.29

3.3. Overtaking speed

Fig. 7 shows the relation between overtaking speed and the three factors pedestrian walking direction, pedestrian walking position, and presence of oncoming traffic.

From the FT model, the parameters walking direction, walking position, and oncoming vehicle presence influenced speed within the 95 %

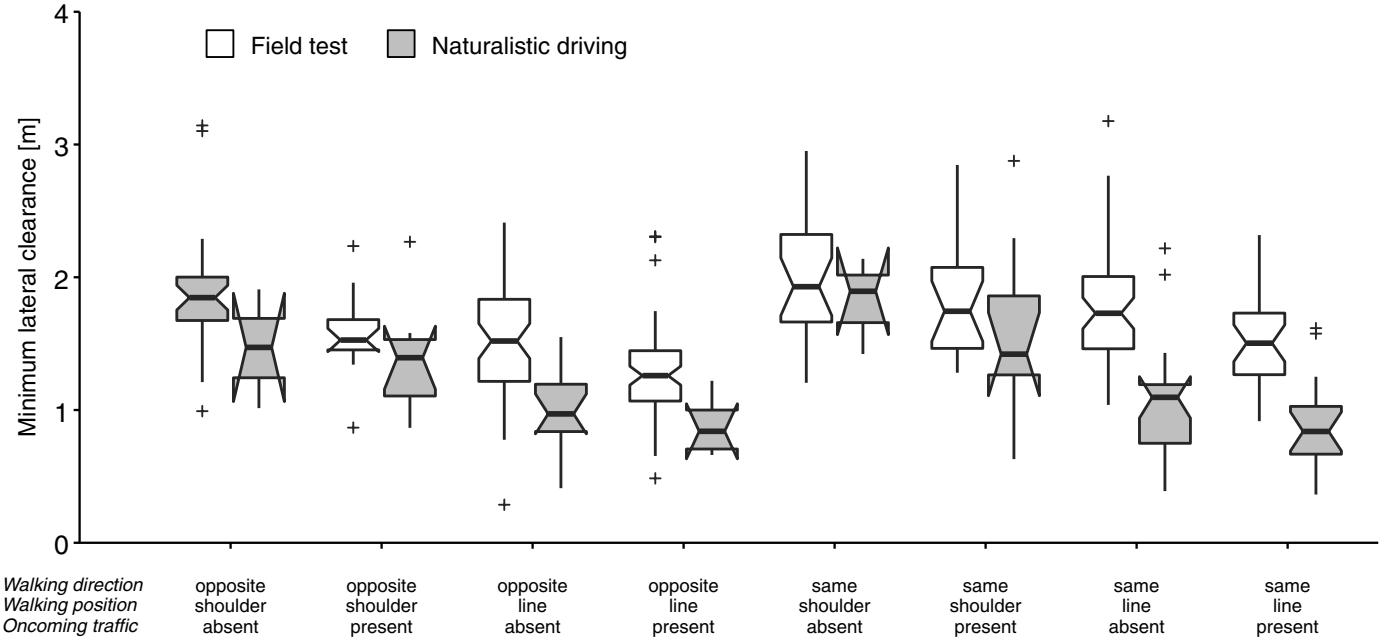

Fig. 6. Minimum lateral clearance for naturalistic driving and field test data, factorized by walking direction, walking position, and oncoming traffic presence. The notches show the +/-1.58/√N interquartile range (IQR, including the number of observations, N) around the medians (thick horizontal line), and the "+"-signs represent outliers. Lower and upper hinges show the 25th and 75th percentiles and vertical lines represent +/-1.5 IQR from the hinges.

Table 5
Median differences in minimum lateral clearance (MLC) between population-level effect outcomes of the model. Lower and upper 95 % highest density intervals (HDIs) are reported as l-95 % HDI and u-95 % HDI, respectively, for naturalistic driving (ND) and field test (FT) data. Each median for ND and FT data is the difference in median between the two effect outcomes.

Factor (difference in outcomes)	ND data [m]	ND model [m]			FT data [m]	FT model [m]		
	Median	Median	l-95 % HDI	u-95 % HDI	Median	Median	l-95 % HDI	u-95 % HDI
Walking direction (opposite – same)	0	0.03	−0.27	0.35	−0.25	−0.28	−0.43	−0.14
Walking position (line – shoulder)	−0.48	−0.59	−1.02	−0.21	−0.33	−0.36	−0.53	−0.19
Oncoming traffic (present – absent)	−0.05	0.07	−0.23	0.39	−0.31	−0.3	−0.44	−0.16

HDI. Like the ND data's model for MLC, the ND data's model for speed did not show any credible trends. The estimates of the full models for overtaking speed are presented in Table A2 in Appendix A.

Table 6 shows the estimates of the optimized speed models for FT and ND data as median values of the location parameter of the Student's t-distribution. LOOCV verified that these reduced models were better fits than the complete models. The expected predictive accuracy was higher by 2.3 (with 1.8 standard deviation) and 2.3 (with 2.8 standard deviation) for FT and ND data, respectively.

Table 7 reports the 95 % HDI of the models and shows the differences between the outcomes of the factors on overtaking speed for ND and FT data and models, respectively, as medians. The validity of our models is supported by the fact that all median differences of all effects for ND and FT data fall within the 95 % HDI of our models. Walking position cannot be described as a significant factor when the ROPE is centered around zero, since the 95 % HDI for the corresponding difference includes zero as well (and hence ROPE and HDI would overlap).

Table 6
Parameter distributions of the optimized models for overtaking speed, fitted from naturalistic driving (ND) and field test (FT) data. Medians are reported together with the lower and upper limits of the 95 % highest density interval (HDI): l-95 % HDI and u-95 % HDI, respectively. $\sigma_{V,ID}$ is the standard deviation of the group-level effect driver identity (ID) and σ_V is the scale parameter of the Student's t-distribution. The number of degrees of freedom of the Student's t-distribution is denoted by ν_V.

Parameter	ND model			FT model		
	Median	l-95 % HDI	u-95 % HDI	Median	l-95 % HDI	u-95 % HDI
β_0	59.87	53.45	66.34	61.07	59.47	62.62
β_{dir}	−3.40	−10.07	3.26	−0.35	−2.26	1.49
β_{pos}	−6.89	−12.05	−1.66	−1.50	−2.87	−0.08
β_{onc}	−5.62	−12.42	0.98	−3.79	−5.70	−1.92
$\sigma_{V,ID}$	2.69	0.00	6.87	N/A	N/A	N/A
σ_V	12.96	10.54	15.52	7.35	6.43	8.30
ν_V	22.57	3.57	50.34	11.64	3.50	24.87

3.4. Time to collision

Fig. 8 shows the TTC at the beginning of the steering away phase for all 48 overtaking maneuvers in the ND dataset without braking. (Of the 83 maneuvers considered for the other measures, 25 with mild braking were excluded, since we assumed that the braking indicated that the driver was aware of the situation and would not have benefited from a warning.) The mean TTC is 3.68 s (median 3.13 s and standard deviation 2.26 s). The dashed line marks the 1.7-s threshold—the latest time at which an FCW system should warn the driver, according to Euro NCAP (Euro NCAP, 2019). Four of the 48 (i.e. 8.3 %) steering maneuvers were performed below this threshold. Walking direction did not influence TTC significantly; in fact, two out of 23 (8.7 %) drivers steered away at TTCs less than 1.7 s when the pedestrian was walking in

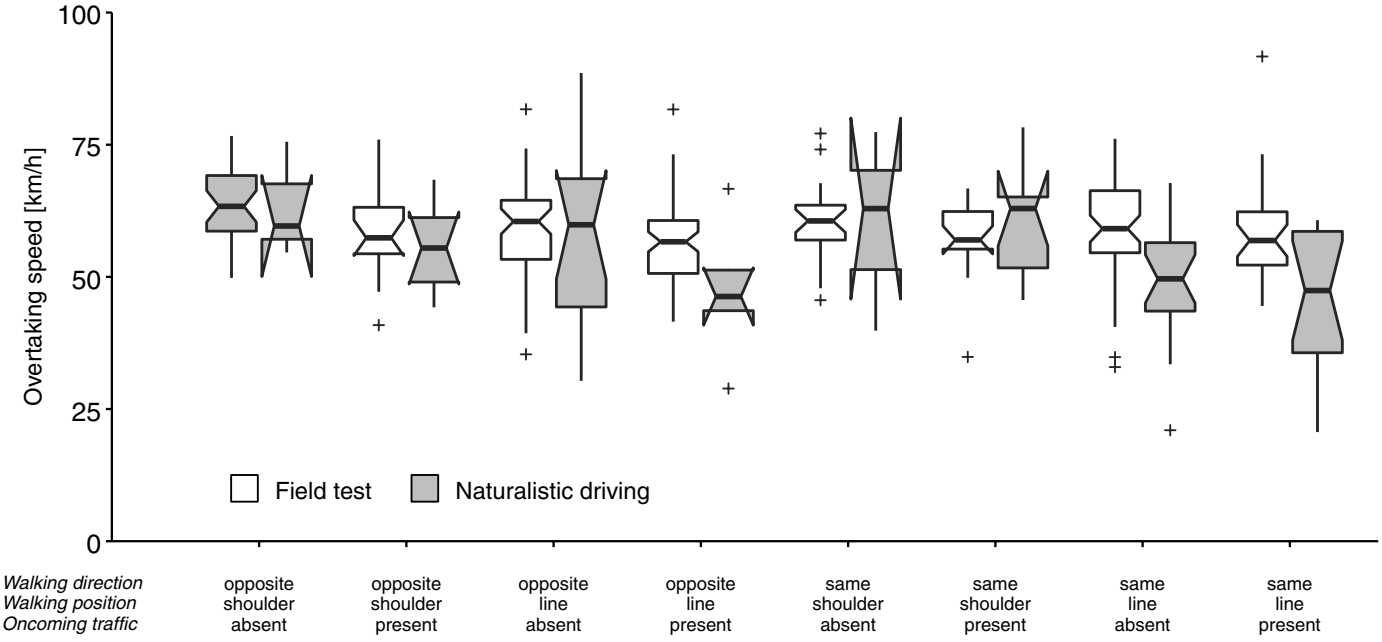

Fig. 7. Overtaking speed for naturalistic driving and field test data, factorized by walking direction, walking position, and oncoming traffic presence. The notches show the +/-1.58/√N interquartile range (IQR, including the number of observations, N) around the medians (thick horizontal line), and the "+"-signs represent outliers. Lower and upper hinges show the 25th and 75th percentiles and vertical lines represent +/-1.5 IQR from the hinges.

Table 7
Median differences in overtaking speed between population-level effect outcomes of the model. Lower and upper 95 % highest density intervals are reported as l-95 % HDI and u-95 % HDI, respectively, for naturalistic driving (ND) and field test (FT) data. The medians for ND and FT data are the differences in medians between the two effect outcomes.

Factor (difference in outcome)	ND data [km/h]	ND model [km/h]			FT data [km/h]	FT model [km/h]		
	Median	Median	l-95 % HDI	u-95 % HDI	Median	Median	l-95 % HDI	u-95 % HDI
Walking direction (opposite – same)	5.8	4.9	−4.2	14.1	−0.3	−0.3	−2.9	2.5
Walking position (line – shoulder)	−8.6	−8.1	−17.8	1.6	−2.6	−2.4	−5.3	0.4
Oncoming traffic (present – absent)	−2.8	−2.9	−11.9	6.1	−4.4	−3.9	−6.5	−1.2

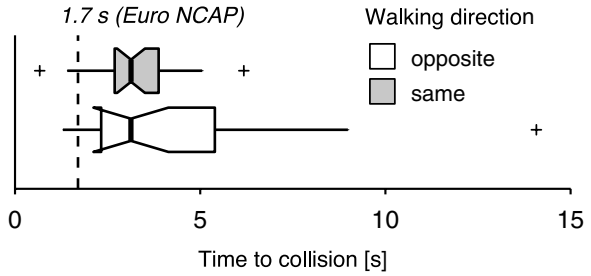

Fig. 8. Time to collision (TTC) at the beginning of the steering away phase for flying overtaking maneuvers in the naturalistic driving dataset. The dashed line marks the 1.7-s threshold on TTC required by Euro NCAP. The notches show the +/-1.58/√N interquartile range (IQR, including the number of observations, N) around the medians (thick horizontal line), and the "+"-signs represent outliers. Lower and upper hinges show the 25th and 75th percentiles and vertical lines represent +/-1.5 IQR from the hinges.

the opposite direction, and two out of 25 (8.0 %) steered away at TTCs less than 1.7 s when the pedestrian was walking in the same direction. The Pearson coefficient for the correlation between TTC and overtaking speed resulted in a small positive correlation (r = 0.20).

4. Discussion

4.1. Driver behavior

Our study has shown that MLC was larger when the pedestrian walked in the direction of the overtaking traffic rather than the opposite direction, while overtaking speed remained unchanged. This result is in line with that reported by Ren et al., who concluded that driver interaction with pedestrians in crossing situations is affected by eye contact—however, with the effect that the safety of the pedestrian is reduced rather than increased (2016). This result also confirms that walking in the opposite direction, as required in traffic regulations (United Nations, 1968), is perceived as safer by drivers. As a result, drivers are comfortable decreasing MLC to pedestrians who are facing them—an unintended behavioral change that can be explained as risk compensation (Jonah et al., 2001; Rudin-Brown and Jamson, 2013; Wilde, 1982).

When pedestrians walked on the edge of the lane (i.e., line scenario), drivers maintained a smaller MLC than when pedestrians walked further from the lane (i.e., shoulder scenario). The difference in the FT data's MLC for the two conditions (median 0.33 m) was, however, smaller than the actual difference between the pedestrian's two positions (0.5 m). Thus, drivers in the shoulder scenario did not deviate from their trajectory as much as in the line scenario. Overtaking speed was lower when pedestrians walked closer to the traffic, confirming that drivers did perceive the situation as less risky when pedestrians walked on the road shoulder. The fact that MLC and speed were less influenced when the pedestrian was further from the lane may, again, be explained in terms of risk compensation (Wilde, 1982)—or simply by the fact that the pedestrian was so far from a possible collision path that the driver was not concerned (Woodworth and Sheehan, 1971).

In the presence of oncoming traffic, drivers drove closer to the pedestrian. The same behavior has previously been found when drivers overtook cyclists in both FT (Dozza et al., 2016) and ND (Feng et al., 2018; Kovaceva et al., 2018) data. This behavior, in line with the literature on risk homeostasis (Damasio, 1995; Näätänen and Summala, 1974; Summala, 1988), can be explained in terms of the field of safe travel being "compromised" by the combined interactions with cyclist/pedestrian and the oncoming traffic (Gibson and Crooks, 1938).

In this study, overtaking speed decreased in the presence of oncoming traffic (albeit significantly only for FT data). In contrast to results reported by Kovaceva et al. (2018) and in line with those of Dozza et al. (2016), no correlation between MLC and overtaking speed was found in this study for either of the datasets. This suggests that drivers' spatial comfort zone does not depend on speed when overtaking pedestrians, although pedestrians may perceive higher speeds as more dangerous. The decrease in variance of MLC, caused by the presence of an oncoming vehicle, hints that drivers reacted more similarly to each other as they were forced towards the boundary of their comfort zone, possibly because of a ceiling effect caused by the criticality of this situation constraining the margins for maneuvering (Yerkes and Dodson, 1908).

4.2. Active safety

Active safety systems may support drivers overtaking pedestrians—for instance, by warning of a potential crash with the pedestrian (FCW) or by actively avoiding it. Active avoidance could mean automatically braking or steering (the latter intervention assumes, of course, that no oncoming traffic is in proximity). The design of these interventions needs to facilitate tuning them to drivers' comfort zone, in order for them to be acceptable (Lübbe, 2015). Models such as the one proposed in this paper may help compute the acceptable thresholds for system interventions and explain how these thresholds depend on factors such as oncoming vehicle presence, pedestrian position, and pedestrian direction. Indeed, a single threshold for systems activation may not fit all drivers, because not all drivers are the same: for instance, some are more aggressive than others (Jonah, 1997). For this reason, Bayesian regression models would be useful for representing an individual driver's behavior in a probabilistic manner; values for MLC and overtaking speed could be sampled from the modeled distributions. These values could be used by an adaptive safety system, which could then customize warnings and interventions to individual drivers. The distribution could be updated after each pedestrian-overtaking maneuver to improve acceptance.

In automated driving, specifically during automated steering maneuvers to overtake pedestrians, our models can inform path planning depending on the factors pedestrian's walking position and direction. By sampling MLC and overtaking speed values from the Bayesian posterior distribution from this study (conditioned on the given scenario), an automated vehicle can perform an overtaking maneuver which is likely to be perceived as comfortable by the driver. According to Abe et al. (2017), a driver's comfort zone should be more conservative in automated driving compared to manual driving, to increase perceived safety. Our Bayesian models can be sampled accordingly to achieve

precautionary safety in automated driving and used in virtual simulations as a driver model (Bärgman et al., 2017).

For the CPLA scenario, Euro NCAP sets a minimum TTC of 1.7 s before which an FCW system must warn a driver (Euro NCAP, 2019). Our study showed that about 92 % of the drivers in the ND dataset steered away from the collision path with the pedestrian before this threshold, independent of the pedestrian's walking direction. Hence, 8% of the drivers would have received a false warning, assuming that they performed the maneuver within their comfort zone.

4.3. Naturalistic and field test data

The trends of MLC from both ND and FT data were similar across conditions, although MLC was generally lower in the former. Most of the results from the ND data (difference in outcomes, Tables 5 and 7) had larger uncertainty than the results from FT data and therefore did not allow to determine significance (Kruschke, 2018). This is probably due to the small size of the ND dataset, the repetitions of drivers, and the fact that these data were confounded by environmental factors which were not present in the FT data. However, our ND data have higher ecological validity than our FT data (Bärgman, 2016). The fact that these two datasets, although collected in very different conditions, showed the same trends proves that combining results from different datasets can provide effective and compelling evidence (Boda et al., 2017).

4.4. Limitations

As noted, one limitation of this study is the intrinsic nature of ND data, which are rife with uncontrolled environmental factors confounding the results. The dearth of significant findings from these data is probably due to the small number of overtaking maneuvers in the ND dataset, as well.

Another limitation is that drivers from two different geographical regions (i.e., France and Sweden), who may have different attitudes to pedestrians, contributed to the ND and FT data, respectively. It can also be argued that individual drivers in France were more exposed to pedestrian-overtaking maneuvers, since some of them accounted for multiple overtaking events in the ND data but each overtaking maneuver in the FT data was probably performed by a different driver.

Furthermore, the measurement setup in the ND data (MobilEye camera system) was different (and possibly less accurate for the estimation of MLC) than the LiDAR used in FT. Hence, the MLC difference between ND and FT data might be attributable to not just a regional difference, but also a difference between the sensors used.

In this study, it was assumed that drivers performed the overtaking while they were within their comfort zone. This might not have been the case under certain conditions and for certain drivers, because driver comfort zone partly depends on driver state (Bärgman, 2016) and personality (Jonah, 1997; Ulleberg, 2001) which were not measured.

4.5. Future work

In this study, we modeled drivers' comfort zone when they were overtaking a pedestrian. However, the comfort of the pedestrian during the maneuver was not investigated. Future studies should also quantify the pedestrian's comfort and behavior. In addition, future research should estimate the latest point (in terms of time or distance) when the collision can still be physically avoided. This information could further support the design of *acceptable* warning and intervention systems that help drivers, as they approach a pedestrian, avoid an uncomfortable situation (for the driver or the pedestrian), or even a collision. Finally, future studies should include driver demographics so that we can understand the extent to which variability in overtaking maneuvers across drivers can be explained by driver characteristics.

5. Conclusions

Our results show that the driver's comfort zone during pedestrian-overtaking maneuvers depends on the walking direction of the pedestrian and the walking position with respect to the lane, as well as on the presence of oncoming traffic, confirming our hypotheses. These results can be explained in terms of risk compensation. Although speed slightly decreased when 1) the pedestrian moved in the same direction, 2) the pedestrian was closer to traffic, and 3) oncoming traffic was present, lateral clearance significantly decreased, suggesting that the risk compensation was subjective (for the driver) and not necessarily objective (for all road users). Therefore, policymakers may use this information to justify and promote regulations on MLC stratified by speed, in order to increase drivers' awareness of the large effect that oncoming traffic has on their behavior and induce them to increase MLC when overtaking pedestrians.

This study is the first to analyze and model the driver's comfort zone when overtaking a pedestrian. We presented a novel methodology for assessing the comfort zone, considering ND data collected from the driver's perspective and FT data from the pedestrian's. Leveraging on the high ecological validity of the ND and the large sample of FT data, this study obtained more solid results than either of the datasets alone would have provided.

The driver comfort zone, as statistically described in this study, could be used to increase drivers' acceptance of the threat assessment and decision making performed by active safety systems. When the activation thresholds for warning systems (e.g., FCW system) or intervention systems (e.g., AEB or AES systems) are set close to the comfort zone boundary, systems' interventions are more likely to be accepted by the drivers.

In the CPLA scenario of Euro NCAP, the current threshold for TTC is 1.7 s. Our results indicate that 8% of the FCW activations using this threshold would be false warnings. This new information could be useful to Euro NCAP as it seeks to design test scenarios for commercial active safety systems that do not exceed the threshold for acceptability.

Bayesian regression modeled the uncertainty in drivers' comfort by giving full posterior distributions of parameters and responses. By incorporating prior knowledge, such models can be used to make active safety systems adaptive to individual drivers and driver states and to support virtual assessment.

CRediT authorship contribution statement

Alexander Rasch: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Gabriele Panero:** Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Christian-Nils Boda:** Methodology, Software, Investigation, Writing - review & editing, Supervision. **Marco Dozza:** Conceptualization, Methodology, Investigation, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

None.

Acknowledgments

The authors would like to thank Alberto Morando for the inspiring discussions and advice about Bayesian regression models. We furthermore thank Jordanka Kovaceva and Erik Svanberg for the support in analyzing the UDRIVE naturalistic driving study. We would like to thank Nils Lubbe for supporting the design of this study and reviewing the manuscript. Finally, we thank Kristina Mayberry for language

revisions of the manuscript.

This study was part of the project *Drivers in Interaction with Vulnerable Road Users*, funded by Toyota Motor Europe, Autoliv, and Veoneer. The work was carried out at the *SAFER Vehicle and Traffic Safety Centre at Chalmers*, Gothenburg, Sweden. Part of the data used in this publication were collected in the UDRIVE project, sponsored by the European Commission.

Appendix A

Table A1
Properties of posterior distribution parameters of the full model for MLC model from naturalistic driving (ND) and field test (FT) data, including median and lower and upper limits of the 95 % highest density interval (HDI).

Parameter	ND data model			FT data model		
	Median	l-95 % HDI	u-95 % HDI	Median	l-95 % HDI	u-95 % HDI
β_0	0.13	−0.15	0.42	0.49	0.44	0.55
β_{dir}	0.15	−0.21	0.49	0.11	0.02	0.20
β_{pos}	−0.28	−0.63	0.08	−0.15	−0.24	−0.07
β_{onc}	−0.04	−0.39	0.30	−0.17	−0.26	−0.08
$\beta_{dir*pos}$	−0.21	−0.71	0.30	0.06	−0.06	0.19
$\beta_{dir*onc}$	−0.13	−0.58	0.33	0.05	−0.08	0.20
$\beta_{pos*onc}$	0.02	−0.47	0.50	0.00	−0.13	0.13
$\beta_{dir*pos*onc}$	0.05	−0.59	0.68	−0.04	−0.23	0.16
$\sigma_{MLC,ID}$	0.20	0.00	0.40	N/A	N/A	N/A
σ_{MLC}	0.39	0.31	0.47	0.27	0.25	0.29

Table A2
Properties of posterior distribution parameters of the full model for overtaking speed model from naturalistic driving (ND) and field test (FT) data, including median and lower and upper limits of the 95 % highest density interval (HDI).

Parameter	ND data model			FT data model		
	Median	l-95 % HDI	u-95 % HDI	Median	l-95 % HDI	u-95 % HDI
β_0	61.10	52.38	70.25	61.61	59.81	63.35
β_{dir}	−6.13	−18.57	6.20	−1.61	−4.21	1.04
β_{pos}	−2.88	−15.20	9.40	−2.96	−5.53	−0.49
β_{onc}	−10.11	−21.68	1.87	−4.60	−7.41	−1.86
$\beta_{dir*pos}$	−5.28	−23.35	11.91	2.89	−0.81	6.64
$\beta_{dir*onc}$	8.34	−7.09	24.55	1.97	−2.07	6.10
$\beta_{pos*onc}$	−2.73	−19.20	13.68	1.42	−2.51	5.34
$\beta_{dir*pos*onc}$	1.05	−21.27	23.06	−1.88	−7.71	3.96
$\sigma_{V,ID}$	2.89	0.00	7.33	N/A	N/A	N/A
σ_V	13.01	10.46	15.62	7.41	6.47	8.35
ν_V	23.38	3.34	51.04	12.38	3.68	26.884

References

Abe, G., Sato, K., Itoh, M., 2018. Driver trust in automated driving systems: the case of overtaking and passing. IEEE Trans. Human-Machine Syst. 48 (1), 85–94. https://doi.org/10.1109/THMS.2017.2781619.

Ambrož, M., 2017. Raspberry Pi as a low-cost data acquisition system for human powered vehicles. Measurement 100, 7–18. https://doi.org/10.1016/j.measurement.2016.12.037.

Bärgman, J., 2016. Methods for Analysis of Naturalistic Driving Data in Driver Behavior Research.

Bärgman, J., Boda, C.-N., Dozza, M., 2017. Counterfactual simulations applied to SHRP2 crashes: the effect of driver behavior models on safety benefit estimations of intelligent safety systems. Accid. Anal. Prev. 102, 165–180. https://doi.org/10.1016/j.aap.2017.03.003.

Barnard, Y., Utesch, F., van Nes, N., Eenink, R., Baumann, M., 2016. The study design of UDRIVE: the naturalistic driving study across Europe for cars, trucks and scooters. Eur. Transp. Res. Rev. 8, 2. https://doi.org/10.1007/s12544-016-0202-z.

Boda, C.-N., 2017. Driver Interaction with Vulnerable Road Users: Understanding and Modelling Driver Behaviour for the Design and Evaluation of Intelligent Safety Systems.

Boda, C.N., Dozza, M., Bohman, K., Thalya, P., Larsson, A., Lubbe, N., 2018. Modelling how drivers respond to a bicyclist crossing their path at an intersection: how do test track and driving simulator compare? Accid. Anal. Prev. 111 (October 2017), 238–250. https://doi.org/10.1016/j.aap.2017.11.032.

Brännstrom, M., Sandblom, F., Hammarstrand, L., 2013. A probabilistic framework for decision-making in collision avoidance systems. IEEE trans. Intell. Transp. Syst. 14 (2), 637–648. https://doi.org/10.1109/TITS.2012.2227474.

Brännström, M., Coelingh, E., Sjöberg, J., 2014. Decision–making on when to brake and when to steer to avoid a collision. Int. J. Veh. Saf. 7 (1), 87. https://doi.org/10.1504/IJVS.2014.058243.

Bürkner, P.-C., 2017. Brms: an r package for bayesian multilevel models using stan. J. Stat. Softw. 80, 1. https://doi.org/10.18637/jss.v080.i01.

Bürkner, P.-C., 2018. Package "brms.". doi:10.18637/jss.v080.i01&. .

Chen, Q., Lin, M., Dai, B., Chen, J., 2015. Typical Pedestrian Accident Scenarios in China and Crash Severity Mitigation by Autonomous Emergency Braking Systems. SAE Techhttps://doi.org/10.4271/2015-01-1464. Pap. 2015-April April.

Damasio, A., 1995. Damasio, Antonio R., Descartes' Error: Emotion, Reason, and the Human Brain, Relations Industrielles. https://doi.org/10.7202/051028ar.

Dozza, M., Schindler, R., Bianchi-Piccinini, G., Karlsson, J., 2016. How do drivers overtake cyclists? Accid. Anal. Prev. https://doi.org/10.1016/j.aap.2015.12.008.

Dozza, M., Rasch, A., Boda, C.N., 2017. An open-source data logger for field cycling

collection: design and evaluation. Int. Cycl. Saf. Conf (September), 6–9.
Euro NCAP, 2018. Euro NCAP 2025 Roadmap.
Euro NCAP, 2019. Test Protocol - AEB VRU Systems.
European Road Safety Observatory, 2016. Traffic Safety Basic Facts 2016 - Pedestrians. https://doi.org/10.1136/bmj.330.7487.367.
Feinberg, F.M., Gonzalez, R., 2012. Bayesian modeling for psychologists: an applied approach. APA Handb. Res. methods Psychol 2 (March 2007), 445–464. https://doi.org/10.1037/13620-024. Vol 2 Res. Des. Quant. Qual. Neuropsychol. Biol.
Feng, F., Bao, S., Hampshire, R.C., Delp, M., 2018. Drivers overtaking bicyclists—an examination using naturalistic driving data. Accid. Anal. Prev. 115 (November 2017), 98–109. https://doi.org/10.1016/j.aap.2018.03.010.
Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with. Communications of the ACM. https://doi.org/10.1145/358669.358692.
Gibson, J.J., Crooks, L.E., 1938. A theoretical field-analysis of automobile-driving. Am. J. Psychol. 51 (3), 453–471.
Hoff, P., Casella, G., Fienberg, S., Olkin, I., 2006. A First Course in Bayesian Statistics. Springerhttps://doi.org/10.1016/j.peva.2007.06.006.
Hoffman, M.D., Gelman, A., 2011. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo 2008. pp. 1–30. https://doi.org/10.1190/1.3627885.
Howard, C., Linder, A., 2014. Review of Swedish Experiences Concerning Analysis of People Injured in Traffic Accidents. pp. 34.
Jonah, B.A., 1997. Sensation seeking and risky driving: a review and synthesis of the literature. Accid. Anal. Prev. 29 (5), 651–665. https://doi.org/10.1016/S0001-4575(97)00017-1.
Jonah, B.A., Thiessen, R., Au-Yeung, E., 2001. Sensation seeking, risky driving and behavioral adaptation. Accid. Anal. Prev. 33 (5), 679–684. https://doi.org/10.1016/S0001-4575(00)00085-3.
Kovaceva, J., Nero, G., Bärgman, J., Dozza, M., 2018. Drivers overtaking cyclists in the real-world: evidence from a naturalistic driving study. Saf. Sci. https://doi.org/10.1016/j.ssci.2018.08.022.
Kruschke, J.K., 2014. Doing Bayesian Data Analysis: a Tutorial With R, JAGS, and Stan, 2nd ed. https://doi.org/10.1016/B978-0-12-405888-0.09999-2. Doing Bayesian Data Analysis: a Tutorial With R, JAGS, and Stan, Second Edition.
Kruschke, J.K., 2018. Rejecting or accepting parameter values in bayesian estimation. Adv. Methods Pract. Psychol. Sci. 1 (2), 270–280. https://doi.org/10.1177/2515245918771304.
Lai, F., Carsten, O., Schmidt, E., Petzoldt, T., Pereira, M., Alonso, M., Perez, O., Utesch, F., Baumann, M., 2013. Study Plan. UDRIVE Deliverable 12.1. EU FP7 Project UDRIVE Consortium. https://doi.org/10.26323/UDRIVE_D12.1.
Laird, J., Page, M., Shen, S., 2013. The value of dedicated cyclist and pedestrian infrastructure on rural roads. Transp. Policy (Oxf) 29, 86–96. https://doi.org/10.1016/j.tranpol.2013.04.004.
Ljung Aust, M., Engström, J., 2011. A conceptual framework for requirement specification and evaluation of active safety functions. Theor. Issues Ergon. Sci. 12 (1), 44–65. https://doi.org/10.1080/14639220903470213.
Lubbe, N., Davidsson, J., 2015. Drivers' comfort boundaries in pedestrian crossings: a study in driver braking characteristics as a function of pedestrian walking speed. Saf. Sci. https://doi.org/10.1016/j.ssci.2015.01.019.
Lubbe, N., Rosén, E., 2014. Pedestrian crossing situations: quantification of comfort boundaries to guide intervention timing. Accid. Anal. Prev. https://doi.org/10.1016/j.aap.2014.05.029.
Lübbe, N., 2015. Integrated Pedestrian Safety Assessment: A Method to Evaluate Combinations of Active and Passive Safety.
Madgwick, S.O.H., 2010. An efficient orientation filter for inertial and inertial/magnetic sensor arrays. x-io. . https://doi.org/10.1109/ICORR.2011.5975346.
Matson, T.M., Forbes, T., 1938. Overtaking and passing requirements as determined from a moving vehicle. Highw. Res. Board Proc. 18, 100–112.
Morando, A., 2019. Drivers' Response to Attentional Demand in Automated Driving.
Näätänen, R., Summala, H., 1974. A model for the role of motivational factors in drivers' decision-making. Accid. Anal. Prev. 6 (3–4), 243–261. https://doi.org/10.1016/0001-4575(74)90003-7.
Quigley, M., Conley, K., P Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R., Y Ng, A., 2009. ROS: An Open-source Robot Operating System. ICRA Workshop on Open Source Software.
Rajamani, R., 2012. Vehicle Dynamics and Control. Springer ISSN 0941 - 5122.
Ren, Z., Jiang, X., Wang, W., 2016. Analysis of the influence of pedestrians' eye contact on drivers' comfort boundary during the crossing conflict. Procedia Eng. 137, 399–406. https://doi.org/10.1016/j.proeng.2016.01.274.
Rudin-Brown, C., Jamson, S., 2013. Behavioural adaptation and road safety: theory. Evidence and Action.
Rusu, R.B., 2010. Semantic 3D object maps for everyday manipulation in human living environments. KI - Künstliche Intelligenz 24 (4), 345–348. https://doi.org/10.1007/s13218-010-0059-6.
Rusu, R.B., Cousins, S., 2011. 3D is here: point cloud library. IEEE Int. Conf. Robot. Autom. 1–4. https://doi.org/10.1109/ICRA.2011.5980567.
Schram, R., Williams, A., Ratingen, Mvan, Ryrberg, S., Sferco, R., 2015. Euro NCAP's first step to assess Autonomous Emergency Braking (AEB) for vulnerable Road users. 24th Int. Tech. Conf. Enhanc. Saf. Veh 1–7.
Summala, H., 1988. Risk control is not risk adjustment: the zero-risk theory of driver behaviour and its implications. Ergonomics 31 (4), 491–506. https://doi.org/10.1080/00140138808966694.
Summala, H., 2007. Towards understanding motivational and emotional factors in driver behaviour: comfort through satisficing. In: Cacciabue, P.C. (Ed.), Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions With Intelligent Transport Systems. Springer, London, London, pp. 189–207. https://doi.org/10.1007/978-1-84628-618-6_11.
Ulleberg, P., 2001. Personality subtypes of young drivers. Relationship to risk-taking preferences, accident involvement, and response to a traffic safety campaign. Transp. Res. Part F Traffic Psychol. Behav. 4 (4), 279–297. https://doi.org/10.1016/S1369-8478(01)00029-8.
United Nations, 1968. 19. Convention on Road Traffic.
Van Nes, N., Christoph, M., Jansen, R., Schagen, I., Verschragen, E., Wesseling, S., et al., 2017. How You Drive: Overview of Main Results of the UDRIVE Project. UDRIVE Overview. EU FP7 Project UDRIVE Consortiumhttps://doi.org/10.26323/UDRIVE_overview.
Van Ratingen, M., Williams, A., Lie, A., Seeck, A., Castaing, P., Kolke, R., Adriaenssens, G., Miller, A., 2016. The european new Car assessment programme: a historical review. Chin. J. Traumatol. - English Ed. 19 (2), 63–69. https://doi.org/10.1016/j.cjtee.2015.11.016.
Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27 (5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4.
Wilde, G.J.S., 1982. The theory of risk homeostasis: implications for safety and health. Risk Anal. 2 (4), 209–225. https://doi.org/10.1111/j.1539-6924.1982.tb01384.x.
Wisch, M., Seiniger, P., Pastor, C., Edwards, M., Visvikis, C., Reeves, C., 2013. Scenarios and Weighting Factors for Pre-crash Assessment of Integrated Pedestrian Safety Systems.
Woodworth, R.S., Sheehan, M.R., 1971. Contemporary Schools of Psychology. Methuen & Co. Ltd, London.
World Health Organization, 2018. Global Status Report on Road Safety.
Yanagisawa, M., Swanson, E.D., Philip, A., Najm, W., 2017. Estimation of Potential Safety Benefits for Pedestrian Crash Avoidance / Mitigation Systems.
Yerkes, R.M., Dodson, J.D., 1908. The relation of strength of stimulus to rapidity of habit-formation. J. Comp. Neurol. Psychol. 18 (5), 459–482. https://doi.org/10.1002/cne.920180503.
Zegeer, C.V., Bushell, M., 2012. Pedestrian crash trends and potential countermeasures from around the world. Accid. Anal. Prev. 44 (1), 3–11. https://doi.org/10.1016/J.AAP.2010.12.007.