

Delft University of Technology

Interaction between pedestrians and automated vehicles

A Wizard of Oz experiment

Rodríguez Palmeiro, Ana; Van der Kint, S.; Vissers, Luuk; Farah, Haneen; de Winter, Joost C.F.; Hagenzieker, Marjan

DOI 10.1016/j.trf.2018.07.020

Publication date 2018

Document Version Final published version

Published in Transportation Research Part F: Traffic Psychology and Behaviour

Citation (APA)

Rodríguez Palmeiro, A., Van der Kint, S., Vissers, L., Farah, H., de Winter, J. C. F., & Hagenzieker, M. (2018). Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment. Transportation Research Part F: Traffic Psychology and Behaviour, 58, 1005-1020. https://doi.org/10.1016/j.trf.2018.07.020

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' – Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public. Contents lists available at ScienceDirect

Transportation Research Part F

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

Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment



^a Department of Transport and Planning, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands ^b SWOV Institute for Road Safety Research, Bezuidenhoutseweg 62, 2594 AW Den Haag, The Netherlands ^c Department BioMechanical Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

ARTICLE INFO

Article history: Received 10 September 2017 Received in revised form 27 July 2018 Accepted 28 July 2018 Available online 16 August 2018

Keywords: Pedestrians Automated vehicles Road safety Perception Gap acceptance

ABSTRACT

Automated vehicles (AVs) will be introduced on public roads in the future, meaning that traditional vehicles and AVs will be sharing the urban space. There is currently little knowledge about the interaction between pedestrians and AVs from the point of view of the pedestrian in a real-life environment. Pedestrians may not know with which type of vehicle they are interacting, potentially leading to stress and altered crossing decisions. For example, pedestrians may show elevated stress and conservative crossing behavior when the AV driver does not make eye contact and performs a non-driving task instead. It is also possible that pedestrians assume that an AV would always yield (leading to short critical gaps). This study aimed to determine pedestrians' crossing decisions when interacting with an AV as compared to when interacting with a traditional vehicle. We performed a study on a closed road section where participants (N = 24) encountered a Wizard of Oz AV and a traditional vehicle in a within-subject design. In the Wizard of Oz setup, a fake 'driver' sat on the driver seat while the vehicle was driven by the passenger by means of a joystick. Twenty scenarios were studied regarding vehicle conditions (traditional vehicle, 'driver' reading a newspaper, inattentive driver in a vehicle with "self-driving" sign on the roof, inattentive driver in a vehicle with "self-driving" signs on the hood and door, attentive driver), vehicle behavior (stopping vs. not stopping), and approach direction (left vs. right). Participants experienced each scenario once, in a randomized order. This allowed assessing the behavior of participants when interacting with AVs for the first time (no previous training or experience). Post-experiment interviews showed that about half of the participants thought that the vehicle was (sometimes) driven automatically. Measurements of the participants' critical gap (i.e., the gap below which the participant will not attempt to begin crossing the street) and self-reported level of stress showed no statistically significant differences between the vehicle conditions. However, results from a post-experiment questionnaire indicated that most participants did perceive differences in vehicle appearance, and reported to have been influenced by these features. Future research could adopt more fine-grained behavioral measures, such as eye tracking, to determine how pedestrians react to AVs. Furthermore, we recommend examining the effectiveness of dynamic AV-topedestrian communication, such as artificial lights and gestures.

© 2018 Elsevier Ltd. All rights reserved.

* Corresponding author.

https://doi.org/10.1016/j.trf.2018.07.020 1369-8478/© 2018 Elsevier Ltd. All rights reserved.







E-mail addresses: a.rodriguezpalmeiro-1@tudelft.nl (A. Rodríguez Palmeiro), sander.van.der.kint@swov.nl (S. van der Kint), luuk.vissers@swov.nl (L. Vissers), h.farah@tudelft.nl (H. Farah), j.c.f.dewinter@tudelft.nl (J.C.F. de Winter), m.p.hagenzieker@tudelft.nl (M. Hagenzieker).

1. Introduction

1.1. Mixed traffic of pedestrians and automated vehicles

In Europe, about 26% of road fatalities concern pedestrians (World Health Organization, 2015). Most of these fatalities occur while pedestrians attempt to cross the road in an urban environment (e.g., European Commission, 2016; SWOV, 2010).

Automated vehicles (AVs) may be expected to reduce these accident rates by replacing error-prone drivers with reliable computers. However, in the coming decades, there will be a situation of mixed traffic with both conventional vehicles and AVs driving on the road, giving rise to uncertainty about safety (Sivak & Schoettle, 2015). Unless fully segregated lanes are created, AVs will also be sharing the roads with vulnerable road users such as pedestrians and cyclists. Before AVs are deployed in traffic, pedestrian safety should be assessed and guaranteed.

Currently, many researchers are concerned with the development of computer algorithms that enable the detection of pedestrians using cameras onboard the AV (Ohn-Bar & Trivedi, 2016). Furthermore, ample research is available on how drivers inside the AV take control, for example when a pedestrian enters the road or when another type of impending hazard occurs (e.g., De Winter, Stanton, Price, & Mistry, 2016; Gold, Damböck, Lorenz, & Bengler, 2013). However, research that is focused on the pedestrians themselves is crucial as well, because pedestrians may alter their behavior in response to AVs.

1.2. Interaction challenges between pedestrians and automated vehicles

A model of situation awareness in dynamic decision making developed by Endsley (1995) can be used to reflect on the factors of relevance during an encounter between pedestrians and AVs (Fig. 1).

As seen in Fig. 1, pedestrians' crossing decisions and crossing behavior depend on situation awareness. More specifically, pedestrians predict the behavior of vehicles (Level 3) based on their perception of vehicle and road features (Level 1) and their comprehension of the situation (Level 2). Vehicle features may include speed and distance (Brewer, Fitzpatrick, Turner, Whitacre, & Lord, 2005; Kadali & Vedagiri, 2013; Yannis, Papadimitriou, & Theofilatos, 2013) as well as cues provided by the driver inside the vehicle, such as eye contact and gestures (Habibovic, Andersson, Nilsson, Malmsten Lundgren, & Nilsson, 2016; Keferböck & Riener, 2015; Kitazaki & Myhre, 2015). Situation awareness, crossing decisions, and crossing behaviors of pedestrians are also influenced by environmental and individual factors (see the top of Fig. 1). Individual factors include preconceptions and trust in AVs (Rothenbücher, Li, Sirkin, Mok, & Ju, 2016) as well as knowledge and expectations about the behavior of road users (AVs) (Houtenbos, 2008). In summary, Fig. 1 illustrates that when pedestrians are able to perceive and understand an approaching vehicle's features and the road situation, they are able to make appropriate predictions regarding the behavior of the vehicle. This, in turn, leads to accurate crossing decisions and safe crossing behavior. If, however, pedestrians have inaccurate perception and comprehension about the behavior of the vehicle, this could lead to wrong predictions, a state of elevated confusion and stress, and unsafe crossing decisions (and see George & Dane, 2016; Starcke & Brand, 2012, indicating that stress is associated with decision making).

There are several ways in which pedestrians may make incorrect crossing decisions when interacting with an AV. First, there could be a problem of perception or comprehension, as pedestrians might be unable to distinguish whether they are interacting with a traditional vehicle or with an AV. One of the drawbacks of AVs is that communication between pedestrians and drivers is



Fig. 1. Model of situation awareness in dynamic decision making, describing the interaction between a pedestrian and an AV (adapted from Endsley, 1995).

not always possible, because the driver of the AV may be performing a non-driving task and therefore is not paying attention to the road (Kitazaki & Myhre, 2015; Rothenbücher et al., 2016). This means that pedestrians may be unable to differentiate between a distracted driver in a traditional vehicle and an AV driver performing a non-driving task, such as reading a newspaper. Second, pedestrians may have misplaced trust in AVs, or incorrect expectations about the behavior of AVs because of a lack of previous experience in traffic situations involving those types of vehicles. For example, if pedestrians believe that the approaching vehicle is driving automatically, they may accept a short gap, because they believe that AVs will yield in all cases. Conversely, it is possible that pedestrians will actually cross with a large gap, because they do not trust the AV's capabilities.

One possible remedy to problems of situation awareness is to make the AVs explicitly recognizable, for example via an external sign stating "self-driving" (Hagenzieker et al., 2016). Although such a sign may alleviate problems of perception and comprehension, a sign may also result in confusion, because the presence of an external sign does not guarantee that AV is driving automatically at that particular moment.

In summary, future AVs may be explicitly recognizable (via signs) or non-recognizable (but still include a distracted driver). Both these situations may be confusing and stressful for pedestrians, and may contribute to altered crossing decisions, as explained in Fig. 1. Thus, it is important to analyze the effect of AVs on pedestrians' gap acceptance, as well as on subjectively experienced stress levels and visual cues that are taken into account in the decision to cross.

1.3. Previous research on pedestrian – automated vehicle interaction

Thus far, a number of studies on the interaction between vulnerable road users and AVs have been conducted (e.g., Blau, 2015; CityMobil2, 2016; Charisi, Habibovic, Andersson, Li, & Evers, 2017; Clamann, Aubert, & Cummings, 2017; Habibovic et al., 2016; Malmsten Lundgren et al., 2017; Rothenbücher et al., 2016).

Clamann et al. (2017) experimentally investigated pedestrians' response times to a Wizard of Oz AV carrying a forwardfacing sign that showed a message (e.g., 'Walk', 'Do not walk'). These authors concluded that legacy behaviors, such as gap distance, are more important for pedestrians in deciding whether to cross the road than the message displayed on the AV. Rothenbücher et al. (2016) also used a Wizard of Oz method to assess the behavior of pedestrians when interacting with an AV without an apparent driver. In a naturalistic driving study, the authors used a car seat costume to conceal the driver, and measured pedestrians' behavior in response to the vehicle. Results showed that pedestrians did notice that the car contained no driver, yet pedestrians were found to be "surprisingly capable of managing this breach of normality without any communication cues" (p. 801).

Lagström & Malmsten Lundgren (2015); (see also Malmsten Lundgren et al., 2017) performed a field study in which 13 participants' willingness to cross was assessed while encountering a traditional vehicle or a Wizard of Oz AV. To create this Wizard of Oz setup, the authors used a dummy steering wheel in a right-hand steered car, whereas the real steering wheel was hidden so that the car appeared to be a left-hand steered to the pedestrians. Different vehicle behaviors (stopping, not stopping) and different driver behaviors (eye contact, phone, newspaper, no driver) were tested. The authors found that a majority of pedestrians who encountered an AV with an inattentive 'driver' or no driver behind the wheel reported an unwillingness to cross. Conversely, all participants would be willing to cross the road when they had eye-contact with the driver. The authors recommended that it may be beneficial to provide information to pedestrians by means of an external vehicle display. Accordingly, the authors also demonstrated a LED strip on the upper part of the windshield, which communicated to the outside traffic whether the AV is in automated driving mode and about to yield (Lagström & Malmsten Lundgren, 2015). Similar AV prototypes have been produced by Nissan (LED display showing "after you") and Daimler (Zebra crossing projected onto the road).

Other studies have used questionnaires or interviews to identify the information needs of vulnerable road users who encounter AVs. Blau (2015) used an online questionnaire from which it was concluded that participants prefer segregated infrastructure when interacting with AVs as compared to when interacting with traditional vehicles. Hagenzieker et al. (2016) carried out a questionnaire study in which participants were shown photos of traditional vehicles and AVs with different external signs ("self-driving") from a cyclist's perspective. The authors found no statistically significant differences between AVs and traditional vehicles regarding the participants' level of certainty as to whether the vehicle had noticed them or would stop for them. Fridman et al. (2017) performed a survey study using 200 participants recruited via crowd-sourcing. The respondents rated 30 vehicle-to-pedestrian display concepts mounted at the front of the car. The authors observed that simple indications (e.g., green pedestrian, green text "walk", red text "don't walk") were most indicative of whether it is safe or not to walk, while certain concepts (e.g., arrows, circles, etc) were more ambiguous. Based on an interview study, Kitazaki and Myhre (2015), provided a list of eight tentative recommendations for communication between pedestrians and AVs. For example, they recommended that autonomous vehicles should identify themselves on the body of the vehicle so that others can form an understanding and trust in AVs. However, at present, there remains a paucity of research regarding AV-pedestrian interactions in real-life settings.

1.4. Aim of this study

As stated above, the interaction between pedestrians and vehicles is based on pedestrians' interpretation of vehicle features (e.g., speed, distance) and communication with the driver. Moreover, road users behave in accordance to how they expect other road users to behave. When AVs are introduced on the roads, communication with drivers might not always be possible, and pedestrians could have incorrect expectations regarding the intention of AVs. Additionally, a mixture of traditional and new AVs might share the roads and pedestrians might be unable to distinguish with which type of vehicle they are interacting. This can create problems of perception and comprehension of the vehicle or traffic situation by the pedestrian, leading to wrong predictions of the vehicle's behavior, stress, and changes in pedestrians' crossing decisions (see Fig. 1). Therefore, the aim of the present study was to analyze the effect of approaching AVs on pedestrians' crossing decisions. Specifically, the following research questions were addressed: (1) Do pedestrians' crossing decisions differ between an encounter with an AV and an encounter with a traditional vehicle?, (2) Do pedestrians' crossing decisions differ as a function of the AV's external features ("self-driving" signs, driver's attentional state)?

An experiment in a real crossing environment was performed in which participants (pedestrians) encountered an approaching vehicle with different occupant behavior (attentive, inattentive, reading a newspaper) and different signage (roof sign "self-driving", hood/door signs "self-driving", no signs). The two conditions with the sign were selected from a photo-based questionnaire study by Hagenzieker et al. (2016). We added a scenario in which the driver was reading a newspaper, and the vehicle was not equipped with any type of external signs (see also Habibovic et al., 2016). The idea of these manipulations was to assess the effects of features/information that may be present in AVs in the future. Participants were not informed in advance about the presence of AVs, to test a possible future situation in which AVs are recently introduced on the roads and no training is provided to road users.

The AV was driven by using a Wizard of Oz setup (Dahlbäck, Jönsson, & Ahrenberg, 1993; Habibovic et al., 2016), meaning that a fake 'driver' sat behind the wheel (e.g., performing non-driving tasks such as reading a newspaper) while the vehicle was actually controlled by the passenger next to the 'driver' by means of a joystick. Pedestrians' crossing decisions were defined in terms of gap acceptance as extracted from video images and GPS recordings. Furthermore, participants were asked to complete questionnaires for assessing the factors taken into account in their decisions as well as their perceived level of stress after interacting with the vehicle.

2. Methods

2.1. Participants

Twenty-four participants were recruited at the Delft University of Technology via posters in different faculties and main campus locations, flyers, individual recruitment, and posts on social media. The number of participants was limited by restrictions in the number of days and time that the road could be closed to other traffic to perform the experiment.

The participants were 9 women and 15 men aged between 19 and 30 years (M = 24.50; SD = 2.95). Participants had different nationalities but it was a requirement that they came from a country where vehicles drive on the right-hand side (steering wheel on the left). Eventually, four participants from countries where vehicles drive on the left-hand side were included because they were used to interact with vehicles that drive on the right-hand side after living in the Netherlands for several years. Other requirements to participate in the experiment were being older than 18 years and having a good command of the English language.

During recruitment, it was explained to the participants that they were going to partake in a study that aimed to analyze the interactions between pedestrians and vehicles, not informing them that Wizard of Oz AVs were involved in the experiment. All participants were rewarded with a gift coupon of ϵ 10 at the end of the field study as a compensation for their collaboration in the study.

The research was approved by the Human Research Ethics Committee of the Delft University of Technology and the SWOV (Institute for Road Safety Research) Ethical Committee, provided that pedestrians would not cross the road. Letting pedestrians step onto the road was also not allowed. Therefore, pedestrians had to indicate their crossing intentions in a different way (i.e., step forward, step backward). All participants provided written informed consent.

2.2. Location and safety procedures

The experiment was conducted at the edge of the Delft University of Technology campus. The road was closed during the experimental sessions to ensure safety. Road blockages were created by placing fences and signs. Moreover, a security guard was present during the experiment to prevent other road users from accessing the road. The road consisted of two driving lanes and two parking lanes. A sidewalk was present on one side of the road; pedestrian crossing facilities were not available (Fig. 2). To ensure the safety of all persons involved in the experiment, the person sitting in the driver's seat could immediately resume control over the vehicle in case of emergency or problems with the joystick. However, such situations did not occur during the experiment.

2.3. Independent variables

The first independent variable was the vehicle's appearance. A single vehicle (a white Toyota Prius) was used to prevent introducing a bias in the participants' behavior due to the vehicle type or color. The vehicle's appearance was set according to one of the following five conditions:



Fig. 2. Experimental location.

- 1. **Traditional vehicle (TV)** (Manually driven No signs Attentive driver): The vehicle was driven manually. The driver was paying attention to the road and kept his hands on the steering wheel (Fig. 3a).
- Non-recognizable automated vehicle (AV) (Joystick driven No signs Newspaper driver): The vehicle was driven by the passenger sitting next to the 'driver' using a joystick connected to a laptop that controlled the vehicle. The 'driver' was holding a newspaper in front of him as if he would be reading it. The vehicle was not equipped with external signs (Fig. 3b).
- 3. Automated vehicle with magnetic signs on the hood and door (AVM) (Joystick driven Hood & door sign Inattentive driver): the vehicle was driven with the joystick by the passenger next to the 'driver'. The 'driver' was intentionally not making eye-contact with the participant (as if he would not be paying attention to the traffic situation) and did not have his hands on the steering wheel. The vehicle was equipped with magnetic signs on the hood and the front door on the pedestrian's side, showing the text "self-driving" (Fig. 3c). The hood magnet measured 80 × 30 cm, and the door magnet measured 60 × 30 cm.
- 4. Automated vehicle with signs on the roof (AVR) (Joystick driven Roof sign Inattentive driver): This was the same as the previous condition, but now the vehicle was equipped with a roof sign with the message "self-driving" (Fig. 3d). The size of that sign was 70×25 cm.
- 5. Traditional vehicle joystick (TVJ): (Joystick driven No signs Attentive driver): Because it was suspected that differences in driving behavior might appear due to the use of the joystick, a fifth scenario was used. In this scenario, the vehicle was driven with the joystick by the passenger but the vehicle looked like a traditional vehicle from the outside (i.e., the 'driver' had his hands on the steering wheel, and he was intentionally making eye-contact with the participants) (Fig. 3a).

It is noted that only 5 types of vehicle appearances out of the 18 possible combinations (i.e., 2 driving modes \times 3 sign conditions \times 3 attention conditions) were tested. For example, we did not test an AV without signs and with an inattentive driver. The reason for this non-crossed design is that the experiment would otherwise take too long per participant.

The second and third independent variables were the vehicle's stopping behavior (stop vs. not stop) and its approach direction (left vs. right). In all scenarios, the vehicle started driving and increased its speed to 25 km/h. This speed was reached at a distance of 70 m relative to the participant. The vehicle then gradually reduced its speed to 10-15 km/h. This speed was reached at a distance of about 20 m from the participant. Finally, the vehicle either stopped or did not stop before the participant. The vehicle approached the participant from either the left or the right. Thus, 20 scenarios (5 appearances × 2 stopping conditions × 2 approach directions) resulted from the three independent variables. The participants experienced each of the 20 scenarios once and in a randomized order.

2.4. Procedures

The experiment took place on four days in March 2017, each day between 09:30 and 15:30. A maximum of seven participants took part per day. The experiment took about 60 min per participant. Different experimenters were involved across



a) Traditional vehicle (TV): Manually driven – No signs – Attentive driver



b) Non-recognizable automated vehicle (AV): Joystick driven – No signs – Newspaper driver



 c) Automated vehicle with magnetic signs on the hood and door (AVM): Joystick driven – Hood & door sign – Inattentive driver.



d) Automated vehicle with signs on the roof (AVR): Joystick driven – Roof sign – Inattentive driver

Fig. 3. Vehicle configurations used in the experiment.



Fig. 4. Top view illustration of the experiment. (1) Start driving at 25 km/h (70 m from the participant); (2) Reduce speed to 10–15 km/h (around 20 m from the participant); (3) Stop/not stop before the participant; (4) Change signs and turn around to start driving on the lane corresponding to the following scenario. Position 1 is about 2 m away from the curb; Position 2 is near the edge of the curb.

the different days, but the experimenters in the car and the experimenter performing the interviews were always the same. The 20 scenarios were randomized across participants. The setup of the experiment in the case of a scenario in which the vehicle approached the participant (pedestrian) from the left is shown in Fig. 4. The dashed black arrows indicate the vehicle's path.

The same protocol was followed for each participant:

- When the participant arrived at the experiment location, one of the experimenters welcomed him/her at the van (Fig. 4). The van was used as an office place because no nearby buildings or facilities were available. The participant signed an informed consent form on which the instructions for the experiment were described.
- Another experimenter walked with the participant to Position 1 on the sidewalk (Fig. 4). There, the experimenter repeated the instructions orally to the participant as follows: You will stand here (Position 1) with your back to the road. When I tell you, you should turn around and walk until here (Position 2). Vehicles with different characteristics will approach you, and you should take one step to the front in the first moment you would cross the road and one step backward in the last moment you would cross the road. It is really important that you do not step onto the road. After that I will tell you to come back here (Position 1) and I will ask you some short questions that I will record.

Participants had to stand in Position 1 with their back towards the road to ensure that they would not know in advance which scenario was going to be tested next. In this way, participants could not see that the researchers were changing the vehicle signs or how the vehicle started driving. Participants walked from Position 1 to Position 2 to simulate a real crossing situation in which pedestrians walk on the pavement while a vehicle is approaching. They were not allowed to cross the road due to ethical restrictions, and to ensure the safety of all people involved in the experiment.

- Each of the 20 scenarios was followed by an oral interview ('post-interaction'). At the end of the experiment, another oral interview ('post-experiment') was performed. Both interviews were audio recorded with a mobile phone.
- Back in the van, the participants completed a digital questionnaire and were provided with a debriefing form and a gift coupon. The debriefing form mentioned the aim of the research ("to analyze the impact of automated vehicles on pedes-trian crossing behavior and perceived safety"), disclosed which variables were measured (e.g., critical gap with the aid of video recording), and stated that different scenarios were studied regarding vehicle recognition. The form did not explain the Wizard of Oz technique, but the experimenter faithfully answered any questions a participant had regarding whether the vehicle was a real AV.

In each interaction, the vehicle started driving and gave a light sign (high beams) to the experimenter when the vehicle was 70 m from the participant. The experimenter told the participant to turn around at this moment.

The post-interaction interview consisted of three questions:

- (1) Which factors did you take into account before making the decision to take a step backward? Participants were permitted to provide multiple answers.
- (2) Did you perceive something different than what you would see when interacting with a car in real life? (In case of an affirmative answer: Did this influence your decision making, and how?)
- (3) How stressed were you on a scale from 0 (not stressed at all) to 10 (extremely stressed)?

The post-experiment interview consisted of the following questions:

- (1) How realistic do you think the setup of this experiment was on a scale from 0 to 10? Do you think it was similar to a crossing situation in real life? (Why/why not?)
- (2) Was it natural for you to take a step forward in the first moment you would cross the road and a step backward in the last moment you would do it? (Why/why not?)
- (3) How do you think the vehicle was driving; do you think the vehicle was driving manually or autonomously?

The post-experiment questionnaire, created in Google forms, was divided into four different parts:

- Part 1: Personal data: Age, gender, and country of origin were asked.
- Part 2: Interaction with the vehicle during the experiment. Participants were asked "Did you realize, <u>before making the decision</u> to cross, that in some scenarios the vehicle was equipped with specific signs on the outside with the message self-driving vehicle?" (yes or no). The subsequent questions were "In case of affirmative answer to the previous question, did the signs on the outside of the vehicle influence your decision to cross?" and "If yes, how did the signs on the side and on the front of the different sign types was clearer to you?": "signs on the side and on the front of the vehicle" or a "sign on the vehicle roof" (accompanied with two exemplary photoshopped pictures). Also, the following questions were asked: "Did you realize, <u>before making the decision to cross</u>, that in some scenarios the driver was performing other tasks than driving?" (yes or no) and "In case of affirmative answer to the previous question, did the fact that the driver was performing other tasks than driving influence your decision to cross?" and "If yes, how did the driver performing other tasks influence your decision to cross?" and "If yes, how did the driver performing other tasks influence your decision to cross?" and "If yes, how did the driver performing other tasks influence your decision to cross?" and "If yes, how did the driver performing other tasks influence your decision to cross?" and "If yes, how did the driver performing other tasks influence your decision to cross?" and "If yes, how did the driver performing other tasks influence your decision to cross?" and "If yes, how did the driver performing other tasks influence your decision to cross?" and "If yes, how did the driver performing other tasks influence your decision to cross?" and "If yes, how did the driver performing other tasks influence your decision to cross?".
- Part 3: Sensation seeking scale and trust in self-driving vehicles. Sensation seeking of participants was assessed using a Brief Sensation Seeking Scale (BSSS-8) (Hoyle, Stephenson, Palmgreen, Lorch, & Donohew, 2002). Participants reported their level of agreement with eight statements, including for example "I would like to explore strange places", using a 5-point

Likert scale from '*Strongly disagree*' to '*Strongly agree*'. The trust in self-driving vehicles was assessed using questions in which participants had to show their agreement with six statements, such as "*I trust the self-driving vehicle to interact safely with pedestrians*" on a 5-point Likert scale. This questionnaire was also used by Hagenzieker et al. (2016).

- Part 4: Feedback and further improvements. The participants were asked the following: "Please write any suggestion or comment you might have to improve the current experiment in the future".

As pointed out in Section 2.1, participants were not informed in advance about the presence of AVs. The questions from the post-experiment interview and the questionnaire were provided at the end of the experiment (and not after each interaction) to prevent that the behavior of participants in the next trial is influenced by the questions (e.g., participants could otherwise start focusing more on the vehicle and try to guess if it is driving autonomously or manually).

2.5. Dependent variables

The dependent variables were (1) critical gap, (2) self-reported stress level, and (3) visual cues taken into account when making the decision to cross the road.

2.5.1. Critical gap

Brewer et al. (2005) defined a critical gap as "the time in seconds below which a pedestrian will not attempt to begin crossing the street. If the available gap is greater than the critical gap, it is assumed the pedestrian will cross, but if the available gap is less than the critical gap, it is assumed that the pedestrian will not cross" (Transportation Research Board, 2000, p. 18–14). In the present study, the critical gap in terms of distance and time was estimated.

As the current study focuses on pedestrian's crossing critical gap, the distance when the pedestrian stepped *backward* was analyzed. The step *forward* was not used as a dependent measure, because participants stepped forward in response to the instructions of the experimenter. In a pilot study, participants reported that taking first a step forward and then a step backward felt more natural to them as compared to a procedure where the participants were facing the road the entire experiment. The following critical gap measures were used.

- (a) <u>Critical gap in distance units obtained from manual video analysis</u>. Three cameras were placed on the opposite side of the road from where the participants were standing during the experiment. The three cameras covered a range of about 30 m to the right and the left side from the pedestrian. Orange cones were placed at the edge of the sidewalk, behind the participant, with a distance of 5 m between them. Critical gaps were obtained through visual inspection of the recorded videos, extracting the distance between the participant and the front of the vehicle at the moment the participant took a step backward.
- (b) <u>Critical gap in distance and time units obtained from GPS.</u> Two cameras were placed inside the vehicle, on top of the dashboard. The view of these two cameras covered the road and the sidewalk (Fig. 5). The moment that the participant stepped back was extracted from these camera images. The distance between the participant and the front of the camera was then calculated based on the GPS recordings on board the vehicle. Similarly, the speed of the vehicle and the critical gap in time units (defined as distance/speed) were calculated.



Fig. 5. View from one of the cameras inside of the vehicle.

We decided to use both video-based and GPS-based measures of the critical gap, because each of these two approaches has its own strengths and weaknesses in terms of human coding reliability, objectivity, and temporal resolution. Thus, the two approaches ought to give the same critical gap distances, but may be subject to different types of error and bias, and therefore complement each other.

2.5.2. Self-reported stress

The perceived level of stress was measured using a one-item scale with values ranging from 0 (*not stressed at all*) to 10 (*extremely stressed*) during the interviews after every interaction between the participant and the vehicle.

2.5.3. Self-reported visual cues

The visual cues that participants took into account were assessed by the answers reported by the participants during the post-interaction interviews, post-experiment interviews, and the digital questionnaire.

2.6. Statistical analyses and missing value treatment

The critical gap data, self-reported stress, and self-reported visual cues were analyzed by means of frequency counts, means, and standard deviations. Differences between the five vehicle conditions were analyzed using a repeated-measures ANOVA, for the left and right approaches separately. The stop and non-stop conditions were averaged per participant (see Section 3.1.3 for justification). Multiple comparisons were conducted using the Tukey-Kramer method.

A parametric ANOVA was used instead of a nonparametric test, because there were no reasons to expect inhomogeneity of variances, data were expected to be approximately normally distributed due to the averaging of two trials per participant (stop and non-stop conditions), and ANOVA is robust to violations of its assumptions (e.g., Blanca, Alarcón, Arnau, Bono, & Bendayan, 2017). However, as a robustness check, and to protect against the possibility of heavily tailed distributions (e.g., floor or ceiling effects in the self-reported stress measure, or early critical gaps scores), the repeated-measures ANOVAs were repeated after rank-transformation of the 480 scores (24 participants \times 20 trials) (Conover & Iman, 1981), see Supplementary materials (Table S1).

For the video-based gaps, there were 19 out of 480 (24 participants \times 20 trials) missing values. For the GPS-based distance and speed gaps, there were 42 out of 480 missing values, and for the time gaps, there were 43 out of 480 missing values. These missing values were trials in which a participant had misunderstood the instructions as she wanted to step onto the road, and trials in which the moment of stepping back was not visible on the video recordings by the three cameras in front of the participant or the camera inside the vehicle. For the GPS-based critical gap time, one missing value concerned a trial in which the vehicle had come to a full stop, and so a gap time could not be computed.

The experiment took place on four days. The weather conditions of the experiment were variable (good weather on the first and second day [n = 13], windy weather on the third day [n = 6], rainy and windy weather on the fourth day [n = 5]). The strong winds led to inaudible recordings of the interviews in some cases. For self-reported stress, 3 out of 24 participants had no data because of inaudible recordings due to wind. These 3 participants were excluded from the analysis of self-reported stress. For the remaining 420 trials (21 participants × 20 trials) there were 13 missing values. Missing values were estimated using Euclidean distance imputation of the nearest-neighbor participant.

3. Results

3.1. Manipulation checks

3.1.1. Self-reported driving mode

First, we examined whether participants believed they were actually interacting with an AV. Participants were asked in the post-experiment interview how they thought the vehicle was driven. Results showed that 8 out of 20 participants with available answers thought the vehicle was driven both manually and autonomously depending on the scenario. One participant thought the vehicle was driven autonomously in all scenarios. A total of 9 participants believed that the vehicle was driven manually in all scenarios, and 2 participants were unsure about the driving mode. The remaining four participants provided no available data due to recording issues (i.e., inaudible recordings due to wind).

3.1.2. Self-reported realism

In the post-experiment interview, the participants were asked to rate the realism of the experiment on a scale from 0 to 10. The mean score was 6.42 (SD = 1.83, N = 19). The most frequently reported justifications were that there was only one approaching car and/or that this car was driving slowly (8 participants), that they were not allowed to step onto the road (3 participants), and that there were no pedestrian crossing facilities (2 participants).

From the 19 participants with available answers, 8 reported in the post-experiment interview that the experimental instructions (i.e., to step forward & backward) were natural to them. The remaining participants stated they were sometimes confused about when to step forward or backward. For example, one participant reported: "The step forward was weird; I think once I kind of forgot that I needed to take a step forward. And the step back was fine".

3.1.3. Verification of approach speed

It is important to verify whether the approach speeds were equivalent between the five vehicle conditions. Fig. 6 shows that the vehicle approached with a higher mean speed in the TV condition than in the four joystick-controlled conditions. A repeated-measures ANOVA indicated a significant effect between the five conditions for the vehicle's speed 30 m in front of the pedestrian (F(4, 92) = 17.26, p < 0.001 for left approaches, F(4, 92) = 18.38, p < 0.001 for right approaches, see Table S2 in Supplementary materials for means and standard deviations). Tukey-Kramer multiple comparisons showed significant differences between the four joystick-controlled conditions, but no significant differences among pairs of the four joystick-controlled conditions. The higher speed for the manually controlled car indicates that the results for this condition should be interpreted separately. For example, if participants in the TV condition have a shorter critical gap than in a joystick-controlled condition, this may be because the vehicle in this condition approached with a higher speed.

Fig. 7 illustrates the speed of the vehicle as a function of the distance relative to the participant, for each scenario and participant. Green markers indicate the moment when the participant stepped back (i.e., the critical gaps). It can be seen that in almost all cases the participant stepped back *before* the vehicle started to slow down. In other words, the vehicle's stopping behavior (i.e., stop & not stop conditions) could not affect the observed critical gaps. We therefore averaged the stop and not stop conditions in the subsequent analyses.

3.2. Effect of vehicle condition on participants' critical gap

Table 1 shows the means and standard deviations of the critical gap distance per vehicle condition and approaching direction. A one-way repeated measures ANOVA of the video-based critical gap distances showed no significant differences between the five vehicle conditions for both left approaches (F(4, 92) = 1.42, p = 0.234) and right approaches (F(4, 92) = 0.47, p = 0.759). For the GPS-based critical gap distances, there was also no statistically significant difference for left approaches (F(4, 92) = 1.10, p = 0.363), but there was for right approaches (F(4, 92) = 2.73, p = 0.034). A Tukey-Kramer multiple comparison showed that the critical gap distance for right approaches was higher for the TV condition than for the AVR and TVJ conditions.

Table 2 shows the vehicle speed and time gap at the moment that the participant stepped back. The effects were all statistically significant (F(4, 92) = 15.15, 11.56, 10.27, 10.19 for the speed left, speed right, time gap left, time gap right measures, respectively, p < 0.001 in all four ANOVAs). A multiple comparison showed that the speed was higher, and the time gap was smaller, for the TV condition than for the four joystick-controlled conditions. However, there were no statistically significant differences in speed or time gap between the four joystick-controlled conditions.

In summary, the vehicle in the TV condition approached with a higher speed, which led to considerably shorter time gaps (Table 2) and which may have contributed to the slightly higher distance gaps (Table 1) as compared to the four joystick controlled conditions. There were no significant differences in gap distance, speed, or gap time between the four joystick-controlled conditions.



Fig. 6. Mean speed versus distance between the vehicle and the participant, for each of the five experimental conditions. If the distance is negative, the car approaches from the right from the participant's perspective.



Fig. 7. Speed versus distance between the car and the participant, for each individual encounter. If the distance is negative, the car approaches from the right from the participant's perspective. If the distance is positive, the car approaches from the left from the participant's perspective. Green markers = speed and distance at the moment the participant stepped back, red lines = the car does not stop, black lines = the car stops. Three step-back events are not visible in this figure (at -74 m, -72 m, & -71 m). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1

Critical gap distance, per vehicle condition and approach direction.

	Video-based distance measurement (m)		GPS-based distance measurement (m)		
	Mean	SD	Mean	SD	
1. TV: Manually driven – No signs – Atte	entive driver				
- Approaching from the left	19.50	7.58	25.16	5.70	
- Approaching from the right	20.82	9.49	26.45	10.15	
2. AV: Joystick driven – No signs – Newspaper driver					
- Approaching from the left	17.23	6.84	23.94	7.79	
- Approaching from the right	20.98	9.20	23.68	8.30	
3. AVM: Joystick driven – Hood & door	sign – Inattentive di	river			
 Approaching from the left 	17.59	5.40	23.65	4.89	
- Approaching from the right	21.64	8.63	23.41	9.32	
4. AVR: Joystick driven – Roof sign – Inattentive driver					
- Approaching from the left	19.66	8.24	25.11	7.03	
- Approaching from the right	19.57	9.62	22.72	7.92	
5. TVJ: Joystick driven – No signs – Attentive driver					
 Approaching from the left 	17.97	6.10	23.45	6.14	
- Approaching from the right	20.87	9.09	22.75	8.22	

3.3. Self-reported level of stress

Table 3 shows that perceived levels of stress were low, with an overall average of 2.7 on a scale from 0 to 10. A repeated measures ANOVA revealed no significant differences for the vehicle approaching from the left (F(4, 80) = 1.12, p = 0.353) and the vehicle approaching from the right (F(4, 80) = 1.64, p = 0.172).

3.4. Visual cues taken into account

3.4.1. Post-experiment questionnaire

In the post-experiment questionnaire, in regards to the question "Did you realize, before making the decision to cross, that in some scenarios the driver was performing other tasks than driving?", 83% of the participants (i.e., 20 of 24) answered "yes". From those 20 participants, 14 (i.e., 70%) reported that the signs influenced their decision-making. There were 7

Table 2

Critical gap speed and time gap, per vehicle condition and approach direction.

	GPS speed measurement (km/h)		GPS time measurement (s)		
	Mean	SD	Mean	SD	
1. TV: Manually driven – No signs – Attentive	driver				
- Approaching from the left	16.23	2.41	5.66	0.71	
- Approaching from the right	16.01	2.46	5.53	1.39	
 Approaching from the right AV: Joystick driven – No signs – Newspaper driv Approaching from the left Approaching from the right AVM: Joystick driven – Hood & door sign – Instit 	r driver				
- Approaching from the left	11.63	3.07	7.55	1.46	
- Approaching from the right	12.20	3.82	6.97	1.46	
3. AVM: Joystick driven – Hood & door sign –	Inattentive driver				
- Approaching from the left	12.13	2.39	7.53	1.80	
- Approaching from the right	11.60	3.42	7.67	1.78	
4. AVR: Joystick driven – Roof sign – Inattenti	ive driver				
- Approaching from the left	12.82	3.28	7.22	1.45	
- Approaching from the right	11.70	4.05	7.06	1.27	
5. TVJ: Joystick driven – No signs – Attentive	driver				
- Approaching from the left	11.62	2.42	7.37	1.46	
- Approaching from the right	11.49	2.84	7.40	1.44	

Table 3

Self-reported stress level, per vehicle condition and approach direction.

	Stress level (0-10)			
	Mean	SD		
1. TV: Manually driven – No signs – Attentive driver				
- Approaching from the left	2.85	1.99		
- Approaching from the right	2.68	2.06		
2. AV: Joystick driven – No signs – Newspaper driver				
- Approaching from the left	2.82	1.86		
- Approaching from the right	2.83	1.98		
3. AVM: Joystick driven – Hood & door sign – Inattentive driv	er			
- Approaching from the left	2.71	1.91		
- Approaching from the right	2.98	1.83		
4. AVR: Joystick driven – Roof sign – Inattentive driver				
- Approaching from the left	2.58	1.93		
- Approaching from the right	2.45	1.93		
5. TVJ: Joystick driven – No signs – Attentive driver				
- Approaching from the left	2.26	1.49		
- Approaching from the right	2.54	1.60		

participants who stated that they were more uncertain/doubtful, less safe, or more careful as compared to a traditional vehicle. Three of 14 participants who reported that the signs influenced their decision making explained that they considered that the vehicle would stop for them and that the situation was safe.

Regarding the question "Which of the different sign types was clearer for you?", 19 out of the 20 participants (95%) who did perceive the signs reported that the roof sign was the clearest.

The post-experiment questionnaire further showed that all the participants who answered "yes" to the question "Did you realize, before making the decision to cross, that in some scenarios the driver was performing other tasks than driving?" (20 out of 24, as stated above) identified the task as reading a newspaper or a map. 19 of the 20 participants (95%) who did see the driver reading a newspaper answered yes to the question "Did the fact that the driver was performing other tasks than driving influence your decision to cross?" From them, 6 participants reported that they took a step backward earlier, 5 participants explained that they were more hesitant or careful to cross, and 3 participants reported that they preferred to have eye contact with the driver. One participant explained that he allowed the vehicle to approach him closer because he assumed that it was a self-driving vehicle that would always yield. In the post-experiment questionnaire, there was no specific question regarding the influence of driver inattentiveness concerning the two sign conditions. However, some participants reported taking into account factors such as eye contact or the fact that the driver did not have his hands on the steering wheel (see Section 3.4.2).

3.4.2. Post-interaction interview

The answers to the question "Which factors did you take into account before making the decision to take a step backward?" were coded into different factors. Out of the 480 interviews held (i.e., 24 participants \times 20 scenarios), there were 81 missing interview recordings (e.g., inaudible recordings), thus yielding a total of 399 interviews. Participants reported a total of 736 factors (note that participants could report multiple factors per interview).

The main factors that participants reported as taken into account before making the decision to take a step backward (i.e., to not cross the road anymore) were: speed of the vehicle (240 out of 736 reported factors) and distance from the vehicle, (204 out of 736). The self-driving signs and the driver reading a newspaper were also considered important factors. In the case of the scenarios involving a vehicle with signs "self-driving" (AVM & AVR), the signs were reported as an important factor 37 out of 306 times. In the scenarios where the driver was reading a newspaper, that fact was reported as an important factor 33 times out of 141 times.

Examples of other factors reported by participants were: eye-contact or no eye-contact with the driver (28 out of 736), seeing the hands of the driver on the steering wheel or not (13 out of 736), weather conditions (8 out of 736), vehicle stopping or not stopping (72 out of 736), small vehicle lateral deviations (caused by the joystick control; 10 out of 736), or comments regarding the approaching direction or road (41 out of 736).

3.5. Self-reported trust in self-driving technology

Table 4 shows the means and standard deviations for each statement of the questionnaire to assess trust in self-driving technology.

3.6. Correlational analyses

Table 5 shows a correlation matrix among the age, gender, mean critical gap variables, as well as mean self-reported stress, trust, and sensation seeking. It can be seen that males accepted shorter gaps, and had a higher sensation seeking score than females, where the latter effect was not statistically significant (Table 5). Furthermore, it can be seen that the gap distance and the vehicle speed were strongly correlated. This can be explained by the fact that the vehicle decelerated upon approach (see Figs. 6 & 7); in other words, if the distance gap is larger, the speed of the vehicle is higher as well.

Question	Mean	SD
 In general, I trust the self-driving vehicle I trust the self-driving vehicle to avoid obstacles I trust the self-driving vehicle when overtaking I trust the self-driving vehicle to keep the right lane I trust the self-driving vehicle to keep the lane abead 	3.67 3.75 3.33 3.92 4 13	0.82 0.94 0.87 0.65 0.80
6. I trust the self-driving vehicle to interact safely with pedestrians	3.25	0.94

 Table 4

 Self-reported trust in self-driving technology.

Note. 1: Strongly disagree; 2: disagree; 3: neither agree nor disagree; 4: agree; 5: strongly agree.

Table 5

Pearson product-moment correlation matrix among selected variables.

		1	2	3	4	5	6	7	8
1	Age (years)								
2	Gender (1 = female, 2 = male)	-0.25							
3	Critical gap distance, video-based (m)	-0.11	-0.55						
4	Critical gap distance, GPS-based (m)	-0.07	-0.49°	0.93					
5	Vehicle speed at critical gap, GPS-based (km/h)	-0.11	-0.56	0.90	0.91				
6	Critical gap time, GPS-based (s)	0.14	-0.36	0.67	0.83	0.50			
7	Self-reported stress (0-10)	-0.05	0.16	-0.29	-0.23	-0.25	-0.02		
8	Trust (1–5)	-0.24	-0.16	-0.10	-0.16	0.03	-0.33	-0.07	
9	Sensation seeking (1–5)	0.00	0.34	-0.22	-0.07	-0.13	0.04	-0.02	-0.05

Note. N = 24, except for variable 4 (N = 23) and variable 7 (N = 21). The means and SDs for the sensation seeking scale are available in the Supplementary materials (Table S3).

 * p < .05 for testing the null hypothesis that the correlation is zero.

p < .01 for testing the null hypothesis that the correlation is zero.

4. Discussion

4.1. Main findings: effect of vehicle condition on gap acceptance and self-reported stress

In the near future, there will be a transition period in which a mixture of automated and traditional vehicles will be present on the roads. AVs will initially be rare, and pedestrians may not have received instructions or training on how to interact with those types of vehicles. As a result, pedestrians may hold incorrect expectations about the intentions of AVs, which could lead to changes in pedestrians' crossing decisions. This research aimed to assess whether pedestrians' crossing decisions differ between an encounter with an AV and an encounter with a traditional vehicle, and between AVs with different features (signs, driver's attentional state). Our results did not reveal significant differences in the critical gap in terms of distance and time between the AV and joystick driven traditional vehicle (TVI) conditions.

Results further showed low self-reported levels of stress, and no significant differences in the critical gap or self-reported stress between the different vehicle conditions. This may be explained by results from the post-interaction interviews: the main factors that pedestrians reported to take into account when making the decision to not cross the road were speed and distance from the vehicle.

Analyses of vehicle speed indicated that the manually driven traditional vehicle was driving at a higher speed than the other four conditions, an unintended effect that is attributed to the different vehicle control methods (i.e., manual vs. joy-stick). The difference in speed between the manually driven traditional vehicle condition and the joystick-controlled conditions were presumably caused by the difficulty to change speed using the joystick, which had to be controlled by small hand movements. This increased speed may have contributed to the slightly higher gap distances for the manual condition (see Tables 1 and 2). We also observed individual differences in gap acceptance, where males accepted shorter gaps than females. The fact that males take more risks than females is in agreement with past research (e.g., Zuckerman, Eysenck, & Eysenck, 1978).

Taken together, our findings confirm previous research (Clamann et al., 2017; Rothenbücher et al., 2016) that individual predispositions and legacy behaviors such as distance or speed are more important for pedestrians in deciding to cross the road than are features related to AVs (e.g., external messages on the vehicle). A wealth of prior research concurs that pedestrians' decision to cross a road is dependent on the distance between the vehicle and them, as well as vehicle speed (e.g., Kadali & Vedagiri, 2013; Oxley, Ihsen, Fildes, Charlton, & Day, 2005; Yannis et al., 2013). A possible explanation is that relative distance and speed have direct implications for the pedestrian's safety, whereas vehicle features such as signs or vehicle size (see Yannis et al., 2013) have only indirect implications. It is important to note that pedestrians might take into account other factors, such as non-verbal communication with the driver or vehicle features, when the vehicle is slowly approaching from a short distance (e.g., when the car is creeping forward in a pedestrian crossing zone).

4.2. Visual features taken into account

Although gap acceptance was not statistically affected by the type of vehicle, most of the participants in the postexperiment questionnaire reported to have perceived the "self-driving" signs and to be influenced by those signs in making the decision to cross the road. Participants reported finding the roof sign clearer than the hood/door signs, which could be because the roof sign is clearly visible from the front, whereas the signs on the hood and door are difficult to perceive.

The majority of participants stated that they felt less safe and more doubtful than when interacting with a traditional vehicle. These feelings of unsafety may be due to the fact that there was no eye-contact with the driver, an important feature that pedestrians report to use before making the decision to cross a road (Keferböck & Riener, 2015). The fact that vehicles with the message "self-driving" are currently not present on roads could be another factor causing pedestrians to feel unsafe or confused. In a supplementary analysis, we found that the participants' critical gap slightly increased with trial number (Fig. S1). Future research, carried out across multiple days, could examine how participants adjust their expectations and decisions as they gain more experience with AVs.

The same may apply to the results obtained when the driver was reading a newspaper. The majority of participants noticed the driver holding a newspaper and reported to be influenced by it. Most of the participants reported that they were more hesitant to cross or that they preferred to have eye contact with the driver. These results are in accordance with Malmsten Lundgren et al. (2017), who concluded that perceived safety of pedestrians is worsened when pedestrians are unable to communicate with the driver as is currently done with traditional vehicles. That particular study also concluded that eye contact with the driver led to a more calmed crossing situation for pedestrians.

4.3. Critical gap and perceived level of stress

The critical gaps in the present study ranged between 5.66 and 7.67 s. These values are lower than the value of 9.4 s which we obtained by using an equation provided by Transportation Research Board (2000; assuming a 'start-up time' of 1.9 s, a walking speed of 1.07 m/s, and a road width of 8 m). A possible reason is that default values were considered in the equation for the pedestrian's walking speed and start-up time. Empirical studies on gap acceptance found that pedestrians' median accepted gap across 11 sites was 6.6 s (Brewer, Fitzpatrick, Whitacre, & Lord, 2006), which is consistent with our results.

4.4. Sensation seeking and trust in self-driving technology

Trust in self-driving technology was on average of 3.67 on a scale from 1 to 5, meaning that participants were neutral to positive towards the technology of AVs. However, trust in that technology was relatively low when interaction with pedestrians was involved in the question. A possible reason for this is that AVs are still not driving on the roads so it is not known how they will behave when interacting with pedestrians. Furthermore, to automatically detect pedestrians is a relatively hard task as compared to automatically maintaining lane (Ohn-Bar & Trivedi, 2016).

No statistically significant correlations were found between sensation seeking and trust in self-driving technology, on the one hand, and critical gap or self-reported stress, on the other. As described above, features related to AVs were not considered to be the most important factors for pedestrians in deciding to cross the road, which could explain why the observed correlations were weak.

4.5. Study limitations

Our study has various limitations. First, only a limited number of participants (24) were involved in the experiment and all of them were young students at the Delft University of Technology. People at a technical university are not representative of the general population; they may have a higher affinity towards technology and better spatial skills than average (Wai, Lubinski, & Benbow, 2009).

Results of the post-experiment interview indicated that about half of the participants believed that the vehicle was driven in an automated mode in some or all of the scenarios. This suggests that the Wizard of Oz method was credible for these participants, but that the other half may have recognized the deception that the car was not driven automatically. However, we did not verify what participants thought of the five vehicle conditions individually. For example, it cannot be ruled out that some participants thought that the traditional vehicle was driven automatically. Future research could adopt a more in-depth post-experiment interview to verify the credibility of the Wizard of Oz technique.

The fact that pedestrians were not permitted to actually cross the road (due to ethical reasons) could have also influenced the results of self-reported stress and critical gaps. Some participants reported during the interviews that they knew they did not have to cross the road, and that there was no real danger when interacting with the vehicle.

Additionally, no pedestrian crossing facilities were available. The fact that only one vehicle was used and no other road users were present made the experiment less realistic, as some participants reported. Furthermore, the use of the joystick to control the car in the Wizard of Oz scenarios involved relatively low approaching speeds of the vehicle. All these factors may explain why the overall realism was only 6.42 on a scale from 0 to 10.

Weather conditions could have also affected the results of the experiment. In particular, windy weather led to missing answers during the interview recordings, and some participants in bad weather conditions reported that they could not always read the signs on the vehicle.

4.6. Recommendations for future research

It is recommended that further studies are performed involving a more heterogeneous participant pool. It is also important to examine the effects of signs on roads with different configurations and in different locations (i.e., intersections, vehicle turning).

Measures could be taken to increase the realism of the experiment. For instance, it would be valuable to include multiple vehicles and multiple types of vulnerable road users. Furthermore, besides reflective signs, dynamic forms of signaling of crossing intention and automation modes (e.g., via lights and sounds) could be explored (see Fridman et al., 2017 for an overview of such concepts). Finally, the fact that participants reported to be influenced by the vehicle's appearance, yet no significant differences were found in critical gaps, calls for further research using more fine-grained behavioral measurements that can expose participants' hesitative behaviors when encountering an AV. Accordingly, future research may use eye-tracking or postural sensors in order to uncover how pedestrians' behavior is affected by an approaching AV.

The present study assessed pedestrians' crossing decisions when interacting with AVs for the first time (without previous experience with them). Future studies may focus on how pedestrians' behavior might change with experience. For example, it could be investigated whether experience leads to behavioral adaptation, whereby pedestrians become more cautious, or conversely, take more risks, when encountering an AV. Furthermore, future research should use a fully-crossed design to test all possible combinations of vehicle appearance, driver behavior, stopping pattern, and driving direction. The present study focused on a realistic selection of 20 scenarios that could occur when VRUs encounter AVs.

Acknowledgments

This work is based on the MSc thesis 'Interaction between pedestrians and Wizard of Oz automated vehicles' performed at the Delft University of Technology. The study has been carried out as part of a larger project at SWOV (Institute for Road Safety Research) entitled "Interactions between cyclists, pedestrians and automated vehicles", with funding from the Dutch Ministry of Infrastructure and the Environment. Ana Rodríguez Palmeiro and Joost de Winter are partially supported by

the research programme VIDI with project number TTW 016.Vidi.178.047 (2018–2022; "How should automated vehicles communicate with other road users?"), which is financed by the Netherlands Organisation for Scientific Research (NWO). The authors would like to thank Edwin Scharp and Peter van Oossanen, technicians at the Transport and Planning depart-

ment of the Delft University of Technology, for their help during the performance of the experiment.

Appendix A. Supplementary material

Supplementary data and scripts are available at: https://doi.org/10.4121/uuid:4343cfa8-c246-4bac-bcfa-e3725aceacf7. Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.1016/j.trf. 2018.07.020.

References

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? Psicothema, 29, 552-557.

Blau, M. A. (2015). Driverless vehicles' potential influence on cyclist and pedestrian facility preferences (Doctoral dissertation). The Ohio State University.

Brewer, M., Fitzpatrick, K., Turner, S. M., Whitacre, J., & Lord, D. (2005). Motorist compliance to staged pedestrians for pedestrian treatments at unsignalized intersections. In Proceedings of the ITE 2005 annual meeting and exhibit compendium of technical papers.

Brewer, M., Fitzpatrick, K., Whitacre, J., & Lord, D. (2006). Exploration of pedestrian gap-acceptance behavior at selected locations. *Transportation Research Record: Journal of the Transportation Research Board*, 1982, 132–140.

Charisi, V., Habibović, A., Andersson, J., Li, J., & Evers, V. (2017). Children's views on identification and intention communication of self-driving vehicles. In Proceedings of the 2017 conference on interaction design and children (pp. 399–404). ACM.

CityMobil2 (2016). Experience and recommendations. Retrieved from <https://web.archive.org/web/20180607141012/http://www.citymobil2.eu/en/upload/Deliverables/PU/CityMobil2%20booklet%20web%20final_17%2011%202016.pdf>.

Clamann, M., Aubert, M., & Cummings, M. L. (2017). Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles. In Proceedings of the transportation research board 96th annual meeting, Washington, DC.

Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician*, 35, 124–129. Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies–Why and how. *Knowledge-Based Systems*, 6, 258–266.

De Winter, J. C. F., Stanton, N. A., Price, J. S., & Mistry, H. (2016). The effects of driving with different levels of unreliable automation on self-reported workload and secondary task performance. International Journal of Vehicle Design, 70, 297-324.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. Human Factors, 37, 32-64.

European Commission (2016). Fatalities at 30 days in EU countries. Retrieved from https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/2014_user.pdf>

Fridman, L., Mehler, B., Xia, L., Yang, Y., Facusse, L. Y., & Reimer, B. (2017). To walk or not to walk: Crowdsourced assessment of external vehicle-topedestrian displays. arXiv preprint arXiv:1707.02698.

George, J. M., & Dane, E. (2016). Affect, emotion, and decision making. Organizational Behavior and Human Decision Processes, 136, 47–55.

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*, 1938–1942.

Habibovic, A., Andersson, J., Nilsson, M., Malmsten Lundgren, V., & Nilsson, J. (2016). Evaluating interactions with non-existing automated vehicles: Three Wizard of Oz approaches. In Proceedings of the 2016 IEEE intelligent vehicles symposium (IV), Gothenburg, Sweden (pp. 32–37).

Hagenzieker, M. P., Van der Kint, S., Vissers, L., Van Schagen, I. N. L. G., De Bruin, J., et al. (2016). Interactions between cyclists and automated vehicles: Results of a photo experiment. In International cycling safety conference 2016, Bologna, Italy.

Houtenbos, M. (2008). Expecting the unexpected; a study of interactive driving behaviour at intersections (PhD thesis Delft University of Technology). SWOV Dissertation series. SWOV Institute for Road Safety Research, Leidschendam.

Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality* and Individual Differences, 32, 401–414.

Kadali, B. R., & Vedagiri, P. (2013). Effect of vehicular lanes on pedestrian gap acceptance behaviour. *Procedia-Social and Behavioral Sciences*, 104, 678–687. Keferböck, F., & Riener, A. (2015). Strategies for negotiation between autonomous vehicles and pedestrians. In S. Wolter, A. Riener, B. Pfleging, S. Geisler, & A. van Laack (Eds.), Mensch und computer 2015, Stuttgart, Germany.

Kitazaki, S., & Myhre, N. J. (2015). Effects of non-verbal communication cues on decisions and confidence of drivers at an uncontrolled intersection. In Proceedings of the 8th international driving symposium on human factors in driver assessment, training, and vehicle design, Salt Lake City, Utah.

Lagström, T., & Malmsten Lundgren, V. (2015). AVIP – Autonomous vehicles' interaction with pedestrians. An investigation of pedestrian-driver communication and development of a vehicle external interface (Master of Science Thesis). Gothenborg, Sweden: Chalmers University of Technology.

Malmsten Lundgren, V., Habibovic, A., Andersson, J., Lagström, T., Nilsson, M., Sirkka, A., et al (2017). Will there be new communication needs when introducing automated vehicles to the urban context? In N. Stanton, S. Landry, G. di Bucchianic, & A. Vallicelli (Eds.), Advances in human aspects of transportation (pp. 485–497). Springer International Publishing.

Ohn-Bar, E., & Trivedi, M. M. (2016). Looking at humans in the age of self-driving and highly automated vehicles. *IEEE Transactions on Intelligent Vehicles*, 1, 90–104.

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. Accident Analysis & Prevention, 37, 962–971.

Rothenbücher, D., Li, J., Sirkin, D., Mok, B., & Ju, W. (2016). Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. In Proceedings of the 2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN), New York, NY (pp. 795–802).

Sivak, M., & Schoettle, B. (2015). Road safety with self-driving vehicles: General limitations and road sharing with conventional vehicles (Technical Report No. UMTRI-2015-2). Ann Arbor: University of Michigan Transportation Research Institute. Available at: ">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequence=1&isAllowed=y>">http://deep

Starcke, K., & Brand, M. (2012). Decision making under stress: A selective review. Neuroscience & Biobehavioral Reviews, 36, 1228–1248.

SWOV (2010). Crossing facilities for cyclists and pedestrians. SWOV-factsheet. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid, Leidschendam. Retrieved from https://www.swov.nl/sites/default/files/publicaties/gearchiveerde-factsheet/uk/fs_crossing_facilities_archived.pdf.

Transportation Research Board (2000). Highway capacity manual HCM (Special report No. 209). Washington, D.C.: National Research Council.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. Journal of Educational Psychology, 101, 817–835.

World Health Organization (2015). Global status report on road safety 2015. Geneva, Switzerland.

Yannis, G., Papadimitriou, E., & Theofilatos, A. (2013). Pedestrian gap acceptance for mid-block street crossing. Transportation Planning and Technology, 36, 450–462.

Zuckerman, M., Eysenck, S. B., & Eysenck, H. J. (1978). Sensation seeking in England and America: Cross-cultural, age, and sex comparisons. Journal of Consulting and Clinical Psychology, 46, 139–149.