Modelling perceived risk and trust in driving automation reacting to merging and braking vehicles

He, Xiaolin; Stapel, Jork; Wang, Meng; Happee, Riender

DOI
10.1016/j.trf.2022.02.016

Publication date
2022

Document Version
Final published version

Published in
Transportation Research Part F: Traffic Psychology and Behaviour

Citation (APA)
He, X., Stapel, J., Wang, M., & Happee, R. (2022). Modelling perceived risk and trust in driving automation reacting to merging and braking vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, *86*, 178-195. https://doi.org/10.1016/j.trf.2022.02.016

Transportation Research Part F: Psychology and Behaviour 86 (2022) 178–195

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

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

Modelling perceived risk and trust in driving automation reacting to merging and braking vehicles

Xiaolin He [a,*], Jork Stapel [a], Meng Wang [b], Riender Happee [a]

[a] *Department of Cognitive Robotics, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, the Netherlands*

[b] *Chair of Traffic Process Automation, "Friedrich List" Faculty of Transport and Traffic Sciences, Technische Universität Dresden, Hettnerstraße 3, 01069 Dresden, Germany*

ARTICLE INFO

Keywords:
Perceived risk
Trust
Automated driving
Prediction models
Pupil dilation
ECG

ABSTRACT

Perceived risk and trust are crucial for user acceptance of driving automation. In this study, we identify important predictors of perceived risk and trust in a driving simulator experiment and develop models through stepwise regression to predict event-based changes in perceived risk and trust. 25 participants were tasked to monitor SAE Level 2 driving automation (ACC + LC) while experiencing merging and hard braking events with varying criticality on a motorway. Perceived risk and trust were rated verbally after each event, and continuous perceived risk, pupil diameter and ECG signals were explored as possible indictors for perceived risk and trust.

The regression models show that relative motion with neighbouring road users accounts for most perceived risk and trust variations, and no difference was found between hard braking with merging and hard braking without merging. Drivers trust the automation more in the second exposure to events. Our models show modest effects of personal characteristics: experienced drivers are less sensitive to risk and trust the automation more, while female participants perceive more risk than males. Perceived risk and trust highly correlate and have similar determinants. Continuous perceived risk accurately reflects participants' verbal post-event rating of perceived risk; the use of brakes is an effective indicator of high perceived risk and low trust, and pupil diameter correlates to perceived risk in the most critical events. The events increased heart rate, but we found no correlation with event criticality. The prediction models and the findings on physiological measures shed light on the event-based dynamics of perceived risk and trust and can guide human-centred automation design to reduce perceived risk and enhance trust.

1. Introduction

Automated vehicles (AV) have the potential to improve safety and comfort and reduce congestion (ERTRAC, 2017). Current production vehicles support SAE Level 1 or Level 2 driving automation systems with Adaptive Cruise Control (ACC) and Lane Centring (LC). Such systems still require continuous supervision by drivers to guarantee safety. Higher automation levels (SAE Level 3 +) gradually allow drivers to shift attention away from dynamic driving tasks but require them to be fallback-ready in case of automation failure (SAE, 2021). Driving automation changes the driver's perceived risk and trust, influencing acceptance of driving automation

* Corresponding author.
E-mail addresses: X.He-2@tudelft.nl (X. He), J.C.J.Stapel@tudelft.nl (J. Stapel), Meng.Wang@tu-dresden.de (M. Wang), R.Happee@tudelft.nl (R. Happee).

https://doi.org/10.1016/j.trf.2022.02.016
Received 20 July 2021; Received in revised form 15 January 2022; Accepted 20 February 2022
Available online 3 March 2022

(Nordhoff, Malmsten, et al., 2021; Zoellick et al., 2019). If the public does not widely accept driving automation, road safety and traffic benefits will not accrue (Dong et al., 2019; Nordhoff, Malmsten, et al., 2021; Noy et al., 2018). Therefore, understanding perceived risk and trust in driving automation is of great importance. Below, we review existing perceived risk and trust studies to synthesize definitions, influencing factors, and measurement methods for the current paper.

1.1. Theories of perceived risk and trust

Perceived risk, or perceived safety, captures the level of risk experienced by users of driving automation. It can differ from operational (or actual) risk (Griffin et al., 2020; Kolekar et al., 2020), which is defined as the combination of accident probability and severity (ISO, 2018a). A low perceived risk leads to feeling relaxed, safe, and comfortable (Osswald et al., 2012; Xu et al., 2018), while a high risk perception results in cautious behaviour (Griffin et al., 2020). Perceived risk is highly individualized and is influenced by personal experience, personality, and attitudes (Jin et al., 2020; Ping et al., 2018). The driving environment, such as urban or rural roads, influences both operational and perceived risk (Cox et al., 2017).

Trust has been studied in psychology, sociology, and human factors. A survey by Kaplan et al. (2020) listed 18 commonly-cited but distinct definitions of trust. Although definitions vary, there are three essential factors in trust: risk of losing or not gaining; uncertain outcomes and interdependence between trustor and trustee. The most cited trust definition is 'the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other party will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party' (Mayer et al., 1995), which is also applicable to driving automation. Subjective thinking, feeling, and emotions account for most trusting behaviour (Lee and See, 2004). The process of trust calibration is dynamic with a high degree of volatility (Lee and See, 2004). Kraus, Scholz, Stiegemeier, et al. (2020) proposed a dynamic trust calibration model based on the theory of Lee and See (2004), and Hoff and Bashir (2015), where initial learned trust is formed before the actual interaction with a system and is related to personality, provided information and driving experience. Dynamic learned trust evolves through interacting with the system and is influenced by experienced performance in similar situations and can be moderated by expectations and presentation (e.g. interface design). Trust factors can be divided into three broad categories: human-related, system-related and environment-related factors (Hancock et al., 2020; Hoff and Bashir, 2015; Kaplan, Kessler, Brill, et al., 2021; Schaefer, 2013). Surveys and interviews demonstrated that trust is affected by personal characteristics (e.g., culture, age, and gender), external situations (e.g., task difficulty, traffic), internal situations (e.g., perceived risk, mood), driver's experience and system's performance. Kraus, Scholz, and Baumann (2020) developed a personality model for trust in automation, indicating that a priori acceptability of automated driving is positively related to trust. Increased trust with experience is reported by Gold et al. (2015) for driving automation and Kaplan, Kessler, Brill, et al. (2021) for artificial intelligence system operators.

Exploring the influencers of perceived risk and trust in driving automation is challenging since most people do not have experience in driving automation, making it difficult to reveal user preferences based on empirics. Recent field experiments showed that real driving automation experience led to improvements in trust (Walker et al., 2018; Xu et al., 2018). However, further work is needed to study the influence of driving scenarios and personal characteristics on perceived risk and trust.

Several studies reported highly negative correlations between perceived risk and trust in driving automation. Some studies regard perceived risk as the antecedent of trust (Li et al., 2019; Nordhoff, Stapel, et al., 2021): where perceived risk is firstly influenced by traffic condition, age, gender, etc., upon which trust is then built. In another school of thought, perceived risk is treated as the descendant of trust (Choi and Ji, 2015; Xu et al., 2018; Zoellick et al., 2019), where trust is firstly formed and then influences perceived risk. The relation between perceived risk and trust is also considered a mutual interaction (Garidis et al., 2020; Pyrialakou et al., 2020). Existing studies include conceptual models, surveys and post-experiment questionnaires. However, the relation between perceived risk and trust still needs further investigation using actual driving automation.

1.2. Measurements of perceived risk and trust

Questionnaires are widely used to measure perceived risk, trust and other psychological constructs (Schaefer, 2013; Stapel et al., 2019; Xu et al., 2018; Yagoda and Gillan, 2012). Post-test questionnaires mainly reflect recent experience, and responses may be biased towards social norms. Besides, questionnaires increase participants' mental workload and are not suitable for real-time measurement. To overcome these drawbacks, continuous measurement devices are considered. Researchers have used handset controls (Beggiato et al., 2019; Rossner and Bullinger, 2019), sliders (Saffarian et al., 2012), rotary bars (Cleij et al., 2018) and angle sensors within the steering wheel (Kolekar et al., 2020) for the continuous rating of perceived risk, trust and other subjective items.

Behaviour indicators (e.g., braking) have the potential to reflect automation reliance and compliance. Tenhundfeld et al. (2020) used intervention by braking as an indicator of distrust in automated parking in a Tesla Model X. Naturalistic braking profiles have also been used to cluster near-miss events to different risk levels (Xiong et al., 2019).

Physiological indicators (heart activity, skin response, etc.) are widely studied for non-intrusive continuous state assessment. Taylor (1964) measured skin conductance (galvanic skin response; GSR) as an index of perceived risk in various traffic situations. GSR rates were 50 times higher during driving than during quiet sitting but were not correlated with actual traffic conditions (e.g. day-time off-peak, night off-peak). Morris et al. (2017) compared driver's trust to GSR in different automation driving modes, finding lower skin conductance in safe driving modes. Ajenaghughrure et al. (2021) identify brain activity (EEG) and gaze as the most robust indictors of trust in driving automation.

Pupil diameter is a practical tool to investigate perceived risk and trust in AV. For perceived risk, Tang et al. (2018) showed that the change rate of pupil diameter is significantly higher in severe crashes than in minor crashes. For trust, Perello-March et al. (2019)

proposed to use increased and decreased pupil size as potential indicators to classify users' distrust and appropriate trust in driving automation. Therefore, we expect pupil diameter to reflect perceived risk and trust in driving automation.

1.3. Modelling perceived risk and trust

Following the principle of human-centred design, modelling perceived risk in AV has gained attention. Varotto et al. (2018) investigated perceived risk in full range ACC. They proposed a decision model where the driver would choose to deactivate ACC if the perceived risk becomes unacceptable based on the risk allostasis theory (Fuller, 2011). Kolekar et al. (2020) presented the driving risk field capturing perceived risk continuously in the spatial and time domain based on driver's verbal ratings and steer response to obstacles in manual simulator driving. Surrogate Measures of Safety (SMoS) evaluate operational risk in terms of event criticality and can provide a basis to estimate perceived risk. Minimum time to collision (TTC) can show the driver's acceptance threshold of perceived risk when they take actions (e.g. braking) (Kiefer et al., 2005). The inverse of TTC represents the relative visual expansion of the obstacle, which is referred to as looming (Lee, 1976). Hence, SMoS reflect event criticality and the related perceived risk, but this relationship needs experimental support.

Trust in AV has been captured with conceptual models with various structures (Choi and Ji, 2015; Lee and See, 2004; Mayer et al., 1995). Marsh and Dibben (2003) identified three layers of trust in information science: dispositional, situational and learned trust. Hoff and Bashir (2015) extended this to trust in AV and considered different personal characteristics (e.g. culture, personality traits and mood) in the three layers. Empirical models have also been developed. Kraus, Scholz, Stiegemeier, et al. (2020) investigated the dynamic process of trust calibration in partial driving automation and high driving automation, demonstrating that trust increased along with knowledge accumulation. Tenhundfeld et al. (2020) reported that drivers trust automation more and intervene less frequently when using driving automation more. Hu and Wang (2021) proposed a prescribed-performance control barrier function with a dynamic model of trust in ACC, where the human will hand over the control to ACC if the system's performance reaches a certain threshold.

1.4. Objectives

Based on the discussions above, we identified the following research gaps. Firstly, a quantitative model between event criticality and perceived risk and trust is still lacking. Secondly, personal characteristics (e.g. age, driving experience) influence perceived risk and trust, but their impact has not been fully quantified in dynamic driving. Thirdly, the relations between perceived risk and trust are primarily derived from surveys and post-experiment questionnaires but hardly investigated in dynamic driving. Lastly, trust in automation has been studied mainly through surveys regarding higher automation levels (Level 3 and higher), whereas perceived risk has been mainly studied for manual driving. Hence it is unclear whether perceived risk and trust operate differently in Level 2 driving automation.

This study contributes to two main objectives: **Objective 1:** to model perceived risk and trust in SAE Level 2 automation. **Objective 2:** to quantify the impact of personal characteristics by verifying their contributions in the perceived risk and trust models.

We conduct a simulator study with partial automation (SAE Level 2) motorway driving with drivers continuously monitoring the automation. This allows continuous measurement of perceived risk during dynamic interactions with other road users. We focus on aggressive merging (cut in) and hard braking as safety-critical events (Dreger et al., 2020). Questionnaires, continuous measures of perceived risk (e.g. sensors for hands), physiological measures (e.g. GSR, ECG), behaviour indicators (e.g. braking behaviour) and eye behaviour (e.g. pupil diameter) are jointly evaluated to assess their ability to quantify perceived risk and trust in automation.

We develop perceived risk and trust models that include the factors mentioned above to explain and predict perceived risk and trust.

The remainder of this paper is structured as follows. The methods, including experiment design and measures of perceived risk and trust, are introduced in Section 2. The results in Section 3 are followed by a discussion in Section 4 and conclusions in Section 5.

2. Methods

2.1. Participants

Twenty-five participants with at least 3 years of driving experience were recruited. A recruitment advertisement was distributed via email to university employees and students, and advertised on the neighbourhood app NEXTDOOR to citizens living in Delft. 25 participants (6 females and 19 males) joined the experiment. The age ranged from 24 to 76 years, with a mean of 40.6 years (SD = 16.3). Years with a driving license varied from 3 to 55 years (Mean = 19.2, SD = 15.0). 16 of the 25 participants reported no experience in adaptive cruise control (ACC) or lane centring systems (LC).

2.2. Apparatus

2.2.1. Driving simulator

The experiment was conducted at Delft University of Technology on a driving simulator named DAVSi with Yaris cockpit (Fig. 1). In this experiment, the motion platform was not actuated. The environment was shown on the cylindrical 180-degree screen using three high-quality projectors (Khusro et al., 2020). CarMaker 8.0.1 was used to create the motorway traffic environment. A model of an Auris

4 was used to simulate the subject vehicle dynamics. Subject vehicle dynamics and traffic were controlled using Simulink on a real-time simulation system (dSPACE SCALEXIO). Motion data of the subject vehicle and other vehicles were logged at 10 Hz.

Automated lateral control of the subject vehicle was performed by the IPG driver model provided by CarMaker. A non-linear full-range ACC algorithm was used with the following key parameters: $t_d = 1.2$ s (desired time gap to the vehicle in front); $s_0 = 6$ m (minimum space gap at a standstill); $v_0 = 27.78$ m/s (100 km/h, desired velocity when there is no vehicle detected in front) (Mullakkal-Babu et al., 2016).

An indicator on the dashboard was used as a basic HMI displaying the automation's working status with two colours. Green indicated that the system was activated and worked well. Yellow indicated that the driver had to take over control, but this never happened during this experiment.

2.2.2. Questionnaires

A pre-questionnaire collected personal characteristics such as gender, age, years licensed, and prior automation experience (see Appendix A). Initial learned trust and dynamic learned trust were assessed before and after the simulator drive using the questionnaire in Table 1, including related questions on the willingness to hand over control, the need to monitor automation and the willingness to do other activities.

2.2.3. Physiological measures

Three physiological signals were measured to assess their predictive value in monitoring trust and perceived risk: Cardiovascular activity (ECG), galvanic skin response (GSR) and pupil dilation.

ECG was measured on Lead II (between the left inner ankle and right inner wrist, with the ground on the right inner ankle) and recorded using a TMSi amplifier at 1024 Hz (Fig. 1C). Heartbeats were identified using BioSigKit (Sedghamiz, 2018). The MTEO_QRST algorithm was found to produce the most reliable detection (Sedghamiz and Santonocito, 2015). Peak detections were inspected manually for mislabelling and ectopic beats. The resulting detections were then converted to the rate and variability metrics Heart rate (Beats per minute, BPM), Inter-beat interval (IBI), Root mean square of successive inter-beat interval differences (RMSSD), and power in the High-frequency band (HF; 0.15–0.40 Hz) using the heart rate analysis toolkit heartpy 1.2.6. For calculating HF, IBI was re-sampled using 3rd order univariate spline interpolation. Metrics were calculated for "ultra-short-term" windows of 30 s, in which RMSSD and HF variability metrics are acceptable surrogates for 5-minute recordings, according to Baek et al. (2015). Samples were deemed too short to inspect the Low-frequency band (LF; 0.04–0.15 Hz). An increase in RMSSD and HF may indicate increased activity of the parasympathetic nervous system (and hence a state of ease), while a reduction could indicate increased anxiety, but HF is also influenced by breathing (Shaffer and Ginsberg, 2017).

GSR was measured on the right palm with a Groove GSR sensor at 60 Hz and de-convolved into phasic and tonic components using Ledalab-349 (Benedek and Kaernbach, 2010).

Pupil dilation (diameter) was measured at 50 Hz using a Tobii head-mounted eye tracker (Fig. 1D) measuring the left eye and post-processed with a 4 Hz low-pass filter (Kret and Sjak-Shie, 2019).

2.2.4. Verbal ratings of perceived risk and trust and continuous ratings of perceived risk

The experimenter asked two questions during the simulator drive after each event (see the detailed experiment design below). The two questions were "*How dangerous do you think was the previous event*?" and *"To what extent do you trust the driving automation according to the previous performance of the system?"* Meanwhile, participants continuously rated their perceived risk with a pressure sensor fixed on the steering wheel (Fig. 1A), obtaining visual feedback through a LED bar (Fig. 1B). The participants were tasked to press the sensor harder whenever they felt unsafe, where no force (zero active LED) indicated no risk and the maximum (ten active LEDs) meant very high risk. The sensor's scope and sensitivity were calibrated based on the data from Astin (1999) to have a better experience on the ratings and visual feedback of the LED bar. The continuous rating was recorded at 60 Hz.

Fig. 1. Experimental setup. Left side: Driving simulator (DAVSi) at Delft University of Technology. Middle: Participant with all measurement devices. Right side: (A) Pressure sensor for reporting continuous perceived risk. (B) LED bar- visual feedback of reported continuous perceived risk. (C) ECG device TMSi to measure cardiovascular activity. (D) Eye tracker Tobii pro 2 to measure pupil dilation.

Table 1
Trust related questionnaire used before and after the simulator drive (scaled between 1 and 10).

Item	Results before the simulator drive (Mean + Std)	Results after the simulator drive (Mean + Std)	*t*	*p*
To what extent do you trust the described driving automation system? (adapted from Meyer-Waarden and Cloarec, 2022)	6.84 (1.57)	7.92 (1.50)	−2.49	0.016
To what extent are you willing to hand over control to the described automation system? (Self-developed)	6.60 (1.98)	6.56 (2.48)	0.06	0.950
To what extent do you think it is necessary to monitor the described automation system? (adapted from Nordhoff, Stapel, et al., 2021)	2.96 (2.62)	4.56 (2.41)	−2.24	0.029
To what extent are you willing to do other activities (e.g., eating, drinking, checking the phone) while using the described automation system? (adapted from Xu et al. 2018; Gold et al. 2015)	4.40 (2.67)	4.60 (2.57)	−0.27	0.789

2.3. Experiment design and scenarios

Level 2 automation is mainly developed for motorway driving (Huang et al., 2018). Car following (braking) and lane change (merging) account for most driving scenarios on motorways (Zhao et al., 2017) and are used as safety-critical events in a simulator study (Sharma et al., 2020). Therefore, merging by an adjacent vehicle and hard braking by a lead vehicle were selected for our research. The driving automation maintained longitudinal velocity or kept the predesigned distance to the lead vehicle. The reference velocity of the subject vehicle and the traffic vehicles was set to 100 km/h (ISO, 2018b).

The participants monitored the automation driving at the right lane (see Fig. 2 and Fig. 3). A merging vehicle entered the motorway from an on-ramp, passed the subject vehicle, and merged between the subject and lead vehicles. Detection of this merging manoeuvre was implemented as the moment when the centre of the merging vehicle crossed the line. This somewhat late detection was seen as representative of current systems. After this detection, the subject vehicle automation followed the merging vehicle instead of the original lead vehicle. At this exact moment, the original leading vehicle braked strongly to 60 km/h, followed by acceleration to 100 km/h. The merging vehicle braked and accelerated accordingly keeping a safe distance. The initial merging distance and braking intensity were both varied threefold, creating 9 merging with hard braking (MB) events with different criticalities (see Table 2). In addition, a hard braking (HB) event without merging was designed to investigate whether perceived risk and trust differ between hard braking after merging and normal hard braking (Table 2). All 10 events in Table 2 were repeated twice, resulting in 20 events. Hereafter, the event names with ‘a’ or ‘b’ mean the first and second exposure to the event. All events occurred in a single drive through a series of ramps (20 out of 23 ramps) along the road, as shown in Fig. 3. The order of events was randomized with intervals around 1-min between merging locations.

Participants could overrule the automation by using braking or gas pedals whenever they felt this was necessary. However, manual steering was not allowed during the simulator drive. No accidents or automation failures were designed.

2.4. Procedure

The participants were asked to read an information letter about the experiment and sign an informed consent form. The experimenter introduced the driving automation system and the procedure. Then the participants were asked to fill the trust questionnaire (Table 1).

Electrodes were attached to measure GSR and ECG and the eye tracker was installed and calibrated. Participants were seated in the simulator, and sensors were connected to their recording devices. The experimenter introduced the scenarios and the operation of the simulator.

The experimenter trained the participants to use the pressure button to rate perceived risk. Participants used the pressure button by giving a number from 0 to 10 and hold each number for at least 3 s. Subsequently, they had to follow a random number between 1 and 10 provided by the experimenter. In the ensuing practice drive, the participants experienced several merging events. They were asked

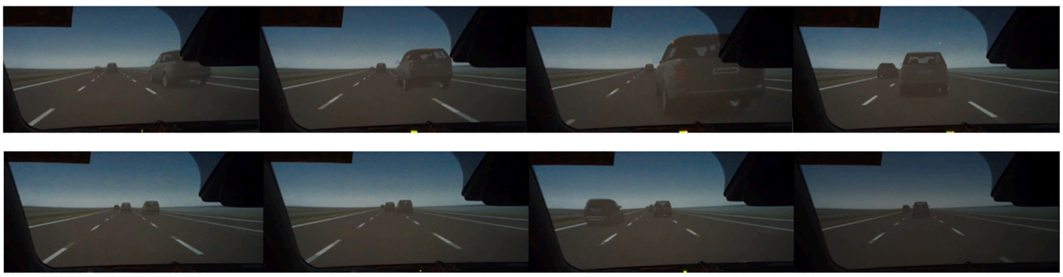

Fig. 2. Video stills of merging with hard braking events. The first row shows the sequence of merging manoeuvre in the most critical MB3 (initial gap 5 m, braking intensity -8 m/s^2), and the second row shows the sequence of the least critical MB7 (initial gap 25 m, braking intensity -2 m/s^2).

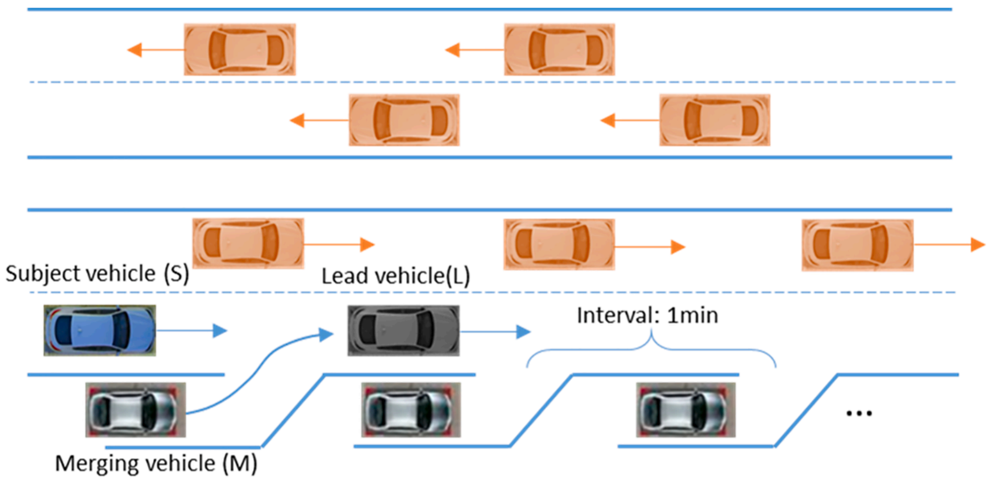

Fig. 3. Merging scenario. The Subject vehicle (S) was driving behind the Lead vehicle (L) when the Merging vehicle (M) merged. L then braked while M and S still had a reduced headway. The scenario's objective risk was varied through merging distance (S-M) and braking intensity of L.

Table 2
Events with different criticalities.

Scenario	Merging gap(initial gap) (m)	Braking intensity (m/s^2)	Event name	minimum THW (s) averaged for the two exposures (measured value)
Merging with hard braking	5	−2	MB1	0.22
		−5	MB2	0.22
		−8	MB3	0.14
	15	−2	MB4	0.48
		−5	MB5	0.48
		−8	MB6	0.48
	25	−2	MB7	0.84
		−5	MB8	0.83
		−8	MB9	0.80
Only hard braking	25	−8	HB1	0.85

to continuously indicate perceived risk using the pressure button and answer the experimenter's questions mentioned in 2.2.4 verbally after each event. The training lasted until the participants could handle all tasks well.

The formal drive (after practice) followed the same procedure, now including 20 events. Another questionnaire with the items in Table 1 was filled after the simulator drive to measure changes in trust.

3. Results

All 25 participants completed the simulator drive, and no motion sickness was reported. 7 out of 25 participants intervened by braking in at least one event. The variations induced by participants' braking and simulated control noise inside the subject vehicle led to 3 events with a close-to-zero distance in the most critical event MB3, one of which was regarded as a collision by the participant. Therefore, these 3 participants were not considered during regression modelling. Besides the three collision participants, 12 outlier ratings (see Fig. 4 for criterion) were removed, leaving 428 out of 500 events for the regression analysis. Pupil diameter and ECG signal were successfully recorded for 22 participants. The recorded GSR signal was not of good quality since skin conductance was outside the device's sensitive range for all but two participants and was excluded from the analysis. We also repeated the experiment without the verbal rating task with 5 extra participants (2 of them are new) to evaluate the influence of speaking on pupil dilation and ECG signals.

3.1. Perceived risk and trust as functions of scenario and personal characteristics

The verbal risk and trust ratings for the events of different criticalities are presented in Fig. 4. Perceived risk varied highly between conditions for all participants. Trust was lowest after the first occurrence of the most critical event (MB3a) and varied less than perceived risk. 12 participants consistently rated trust as 7 or higher in all events except in MB3.

3.1.1. Correlation analysis of potential influencers for perceived risk and trust

Perceived risk and trust are influenced by many factors. Before the regression analysis, we selected four clusters of potential factors (see Table 3). Cluster 1 includes the participant's responses to the events, including verbal ratings of perceived risk (PR) and trust

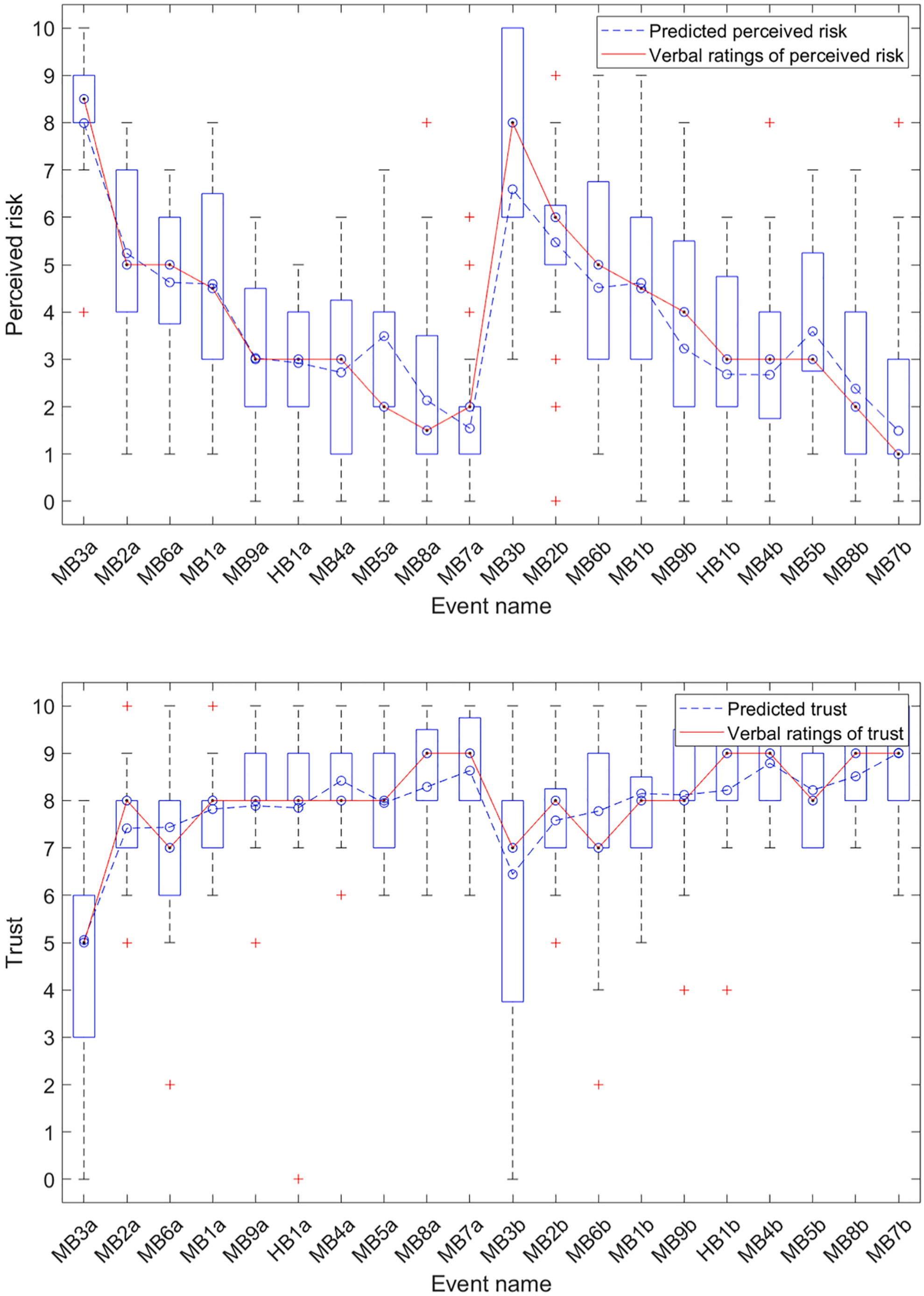

Fig. 4. Upper: Verbal perceived risk ratings for different events. Lower: Verbal trust ratings for different events. Bars present verbal ratings (5, 25, 50, 75, 95 percentile). The dashed blue lines represent the model output in Equation (1) and (2). The red '+' represent outliers beyond 75 percentile + 1.5IQR or below 25 percentile-1.5IQR (IQR is the first quartile subtracted from the third quartile). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(TRU) themselves and the maximum braking by participants (max_B) because they can reflect driver's perceived risk and trust (Beggiato et al., 2015; Cleij et al., 2018; Tenhundfeld et al., 2020; Xu et al., 2018). Cluster 2 captures the criticality of the event with factors related to relative motion, including the initial merging gap (IMG), minimum gap (min_gap), minimum time to collision (min_TTC), minimum time headway (min_THW) and braking intensity (BI) of the merging vehicle during the events. Here min_gap, min_TTC and min_THW express the smallest gap in different manners and are established SMoS to assess the criticality of vehicle interactions (Wang et al., 2021). These factors (e.g., minimum gap, minimum TTC, minimum THW, etc.) have been used to reflect perceived risk in prior studies (Kiefer et al., 2005; Ma et al., 2018). We will also verify their substitution in the regression. Additionally, we will investigate the overlap among influencers of perceived risk and trust. Cluster 3 includes personal characteristics, age (AGE), gender (GEN), years with a driving license (YDL), and automation experience (AE). These factors play essential roles in individual modelling of perceived risk and trust (Gold et al., 2015; Jin et al., 2020; Kaplan, Kessler, Brill, et al., 2021; Kraus, Scholz, Messner, et al., 2020; Ping et al., 2018; Rhodes and Pivik, 2011). Cluster 4 includes repetition (REP) of the event and event type (ET) being merging or hard braking because trust may change with automation experience according to the trust formation from Lee and See (2004), Hoff and Bashir (2015), and Kraus, Scholz, Stiegemeier, et al. (2020). Also, participants perceived less risk in the second exposure to the same event in our experiment (Fig. 4).

Before the stepwise regression, we checked for multicollinearity among potential influencers.

In Cluster 2, non-linear transformations on the original metrics were explored (Appendix B). We found that the logarithm of min_gap, min_TTC and min_THW have strong linear relationships with the participant-averaged perceived risk and the reciprocal of the min_gap, min_TTC and min_THW have strong linear relationships with trust (see Fig. B. 1 in Appendix B). Therefore, the logarithm and the reciprocal of min_gap, min_TTC and min_THW were used as potential predictors for perceived risk and trust in the regression.

We only found strong ($|r| > 0.7$, shaded in Orange in Table 3) correlations within Cluster 2 and Cluster 3. In Cluster 2, strong correlations exist between min_gap, min_TTC, min_THW and IMG (correlation group 1). In Cluster 3, AGE highly correlates with YDL (correlation group 2). Therefore, in the following regression procedure, we will use at most one variable of each of the two correlation groups as potential predictors to resolve multicollinearity.

3.1.2. Stepwise regression analysis of perceived risk and trust

Stepwise regression is an efficient way to select suitable predictors after the elimination of multicollinearity. We used stepwise multiple regression to model the influence of event criticality and personal characteristics on perceived risk and trust. The consecutive steps of both regression models are shown as the first 4 models in Table 4 and Table 5 for perceived risk and trust respectively. Hence, the two models 4 represent the final models of perceived risk (PR) and trust (TRU) as shown in Equation (1) and (2).

$$PR = 9.384 - 2.473 \cdot log(min\_gap) - 0.038 \cdot YDL - 0.201 \cdot BI + 0.470 \cdot GEN \quad (1)$$

$$TRU = 8.780 - 6.265 \cdot (1/min\_TTC) + 0.125 \cdot BI + 0.016 \cdot YDL + 0.372 \cdot REP \quad (2)$$

According to the results, perceived risk and trust mainly vary with min_gap, min_TTC and BI. A smaller minimum gap and more intense braking lead to higher perceived risk; a larger minimum TTC and more gentle braking cause higher trust. Participants with more driving experience trust the automation more and are less sensitive to risk. Female participants are more susceptible to risk. Participants trust the automation more in the second exposure to the events.

The participant-averaged results were used to validate the two models expressed by Equation (1) and Equation (2). A very good fit was obtained (Fig. 4) with R-square = 0.9379, F (2, 17) = 88.2940 ($p = 0.000$) for perceived risk and adjusted R-square = 0.8643, F (3,

Table 3
Correlation analysis (high correlations in orange and insignificant correlations in grey).

		Cluster 1			Cluster 2								Cluster 3				Cluster 4	
		PR	TRU	max_B	Log (min_gap)	Log (min_TTC)	Log (min_THW)	1/ min_gap	1/ min_TTC	1/ min_THW	IMG	BI	AGE	YDL	GEN	AE	ET	REP
Cluster 1	PR	1.000																
	TRU	-0.607***	1.000															
	max_B	0.278***	-0.449***	1.000														
Cluster 2	Log(min_gap)	-0.632***	0.477***	-0.302***	1.000													
	Log(min_TTC)	-0.611***	0.487***	-0.373***	0.918***	1.000												
	Log(min_THW)	-0.579***	0.416***	-0.248***	0.972***	0.919***	1.000											
	1/min_gap	0.592***	-0.513***	0.337***	-0.914***	-0.818***	-0.857***	1.000										
	1/min_TTC	0.566***	-0.599***	0.690***	-0.781***	-0.829***	-0.736***	0.840***	1.000									
	1/min_THW	0.581***	-0.468***	0.275***	-0.951***	-0.871***	-0.949***	0.962***	0.805***	1.000								
	IMG	-0.519***	0.337***	-0.176***	0.907***	0.881***	0.965***	-0.705***	-0.614***	-0.840***	1.000							
	BI	-0.280***	0.307***	-0.157**	0.144**	0.172***	-0.013	-0.218***	-0.243***	-0.071	-0.090*	1.000						
Cluster 3	AGE	-0.179***	0.112*	0.063	-0.041	0.050	-0.001	0.012	-0.021	-0.013	0.005	-0.017	1.000					
	YDL	-0.186***	0.134**	0.095	-0.035	0.033	0.000	0.002	-0.002	-0.019	0.002	-0.018	0.957***	1.000				
	GEN	-0.051	-0.065	0.020	-0.004	-0.017	-0.004	-0.018	-0.010	-0.014	-0.016	-0.011	0.059	0.154***	1.000***			
	AE	0.063	-0.078	0.115*	-0.028	-0.016	-0.019	0.023	0.054***	0.019	-0.015	0.040	-0.333***	-0.240***	0.091	1.000		
Cluster 4	ET	-0.128*	0.000	-0.064	0.283***	0.282***	0.307***	-0.192***	-0.170***	-0.233***	0.326***	-0.275***	0.016	0.005	-0.049	-0.046	1.000	
	REP	0.033	0.121**	-0.037	0.028	-0.073	0.022	-0.083*	-0.026	-0.062	-0.005	0.022	0.002	0.002	-0.001	0.003	-0.015	1.000

*$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)

Table 4
Multiple regression results of perceived risk. Models follow the stepwise regression (Model 4* is the final perceived risk model). See Table B. 1 in Appendix B for alternative models.

Model	Variables in model	R^2	Adjusted R^2	F	p	Unstandardized B	95% confidence interval for B		Standardized coefficients β	t	p
							Lower bound	Upper bound			
1	Constant	0.400	0.398	283.782	0.000	10.183	9.451	10.916		27.326	0.000
	Log(min_gap)					−2.556	−2.854	−2.258	−0.632	−16.846	0.000
2	Constant	0.443	0.440	169.033	0.000	10.981	10.223	11.738		28.499	0.000
	Log(min_gap)					−2.585	−2.873	−2.297	−0.640	−17.657	0.000
	YDL					−0.035	−0.047	−0.023	−0.208	−5.743	0.000
3	Constant	0.481	0.477	130.777	0.000	9.673	8.806	10.540		21.929	0.000
	Log(min_gap)					−2.472	−2.753	−2.191	−0.612	−17.281	0.000
	YDL					−0.035	−0.047	−0.024	−0.211	−6.012	0.000
	BI					−0.202	−0.274	−0.130	−0.196	−5.538	0.000
4*	Constant	0.487	0.482	100.427	0.000	9.384	8.487	10.280		20.565	0.000
	Log(min_gap)					−2.473	−2.752	−2.193	−0.612	−17.377	0.000
	YDL					−0.038	−0.049	−0.026	−0.223	−6.327	0.000
	BI					−0.201	−0.273	−0.130	−0.195	−5.546	0.000
	GEN					0.470	0.071	0.869	0.082	2.313	0.021

Table 5
Multiple regression results of trust. Models follow the stepwise regression (Model 4* is the final trust model). See Table B. 2 in Appendix B for alternative models.

Model	Variables in model	R^2	Adjusted R^2	F	p	Unstandardized B	95% confidence interval for B		Standardized coefficients β	t	p
							Lower bound	Lower bound			
1	Constant	0.359	0.358	238.883	0.000	8.715	8.536	8.894		95.771	0.000
	1/min_TTC					−6.770	−7.631	−5.909	−0.599	−15.456	0.000
2	Constant	0.387	0.384	134.126	0.000	9.299	8.984	9.615		57.969	0.000
	1/min_TTC					−6.300	−7.169	−5.431	−0.558	−14.247	0.000
	BI					0.125	0.069	0.180	0.171	4.379	0.000
3	Constant	0.405	0.401	96.126	0.000	8.974	8.616	9.331		49.332	0.000
	1/min_TTC					−6.290	−7.147	−5.433	−0.557	−14.425	0.000
	BI					0.126	0.071	0.182	0.174	4.509	0.000
	YDL					0.016	0.007	0.025	0.136	3.627	0.000
4*	Constant	0.416	0.410	75.290	0.000	8.780	8.399	9.161		45.326	0.000
	1/min_TTC					−6.265	−7.115	−5.414	−0.555	−14.476	0.000
	BI					0.125	0.071	0.180	0.172	4.500	0.000
	YDL					0.016	0.007	0.025	0.136	3.651	0.000
	REP					0.372	0.107	0.638	0.102	2.754	0.006

16) = 27.6157 (p = 0.000) for trust. The root mean squared error (RMSE) is 0.4044 for perceived risk and 0.3164 for trust while the Pearson correlation is 0.9812 (p = 0.000) for perceived risk and 0.9554 (p = 0.000) for trust, indicating that the models well predict the participant-averaged perceived risk and trust ratings within events. The group based regression models also describe individual events within individual participants quite well, as shown in Fig. B. 2 and Fig. B. 3 in Appendix B, resulting in RMSE = 1.7810, Pearson coefficient = 0.7252 (p = 0.000) for perceived risk and RMSE = 1.4835, Pearson coefficient = 0.5824 (p = 0.000). Note that the output of these two models can be out of the feasible range [0, 10] because the linear regression models do not constrain the output, but individual data were all within the feasible range. The perceived risk model also well predicted the extra 5 participants' data with R-square = 0.3641, F (3, 86) = 17.0563 (p = 0.000) in RMSE = 2.3727, Pearson coefficient = 0.6562 (p = 0.000) (see Fig. B. 4 in Appendix B).

3.1.3. Individual calibration

We calibrated the two models in Equation (1) and (2) for each participant (Appendix C), where both models were statistically significant for all participants except the trust model for participant 18. Female participants were only sensitive to minimum gap but not braking intensity (the models without BI for the female participants have an average RMSE = 1.4794); participants with more driving automation experience perceived the risk to depend only on the minimum gap.

3.1.4. Substitution of potential predictors

As motivated in Section 3.1.1 and 3.1.2, only the best predictor in each of the two correlation groups was adopted in the regression. Below we verify the substitution of alternative predictors.

In correlation group 1, we replaced min_gap and min_TTC in Equation (1) and (2) with the other three relevant factors. For example, in Equation (1), we replaced log(min_gap) with log(min_TTC). This generated models 5–7 (see Table B. 1 and Table B. 2 in Appendix B). The new models remain significant and the R-square decreases slightly. Hence, these common safety metrics have a similar capability to predict perceived risk and trust, indicating that they can be replaced. However, some predictors become insignificant, such as GEN in perceived risk (see model 9 in Table B. 1 in Appendix B) and REP in trust (see model 9 in Table B. 2 in Appendix B).

In correlation group 2, we replaced YDL with AGE (see model 8 in Table B. 1 and Table B. 2 in Appendix B). The new model 8 is still significant with only a slight decrease in R-square, but GEN is no longer significant. Hence, AGE and YDL have a similar performance in predicting perceived risk and trust and are thereby replaceable.

For other predictors outside the two correlation groups, REP only appears in the trust model (Equation (2)), meaning that participants trust the automation more after the first exposure of the events, which implies that trust will accumulate over time provided that no crash or automation failure occurs. To further validate the effects of automation exposure, we compared the trust levels of the

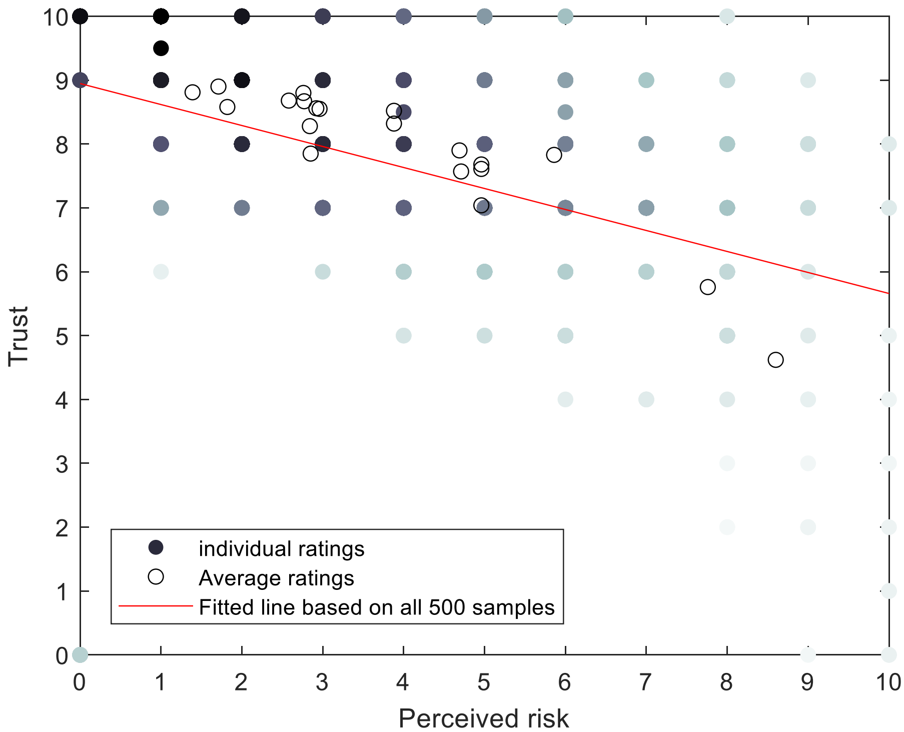

Fig. 5. Relation between perceived risk and trust for individual events (filled dots) and averaged over participants (circles). Darker dots indicate more overlapping data points.

questionnaires before and after the simulator drive for all 25 participants (Table 1). We found only participants 4 and 6 trusted the automation less after the simulator drive, and especially participant 6 reported a 'crash' during the experiment, where the minimum distance between the automated vehicle and the lead vehicle was close to zero. The average trust level increased from 6.84 (SD = 1.54) before the simulator drive to 7.92 (SD = 1.47) ($p = 0.016$) after the drive. Hence, both the inclusion of REP in the trust model and the questionnaire indicate that trust generally increased over time.

3.2. Correlation between perceived risk and trust

Our regression models for perceived risk and trust show a substantial similarity in agreement with the literature (Ayoub et al., 2021; Xu et al., 2018; Zoellick et al., 2019). This similarity is also supported by Fig. 5. A strong linear relationship was found with a Pearson correlation $r = -0.919$ ($p < 0.01$) between participant-averaged perceived risk and trust levels for different events. Regarding all 500 events, the repeated measures Pearson correlation coefficient is -0.649 ($p = 0.000$) (Bakdash and Marusich, 2017). This means that people trust the system more after events where they perceive a lower risk. The individual correlations between perceived risk and trust are significant for 13 out of 25 participants but are not significant for participants with a low standard deviation of trust (Table B. 3 in Appendix B).

To evaluate how well the predictors discriminate between trust and perceived risk, we cross-validated the models of perceived risk and trust. Specifically, we used the predictors of perceived risk to model trust and the predictors of trust to model perceived risk (models 9 in Table B. 1 and Table B. 2 in Appendix B). The two models are still significant with a slightly lower R-Square but REP and GEN are no longer significant in the new models.

We conclude that perceived risk and trust negatively correlate and can be modelled using the same predictors of min_gap, min_TTC, YDL, and BI. However, REP only significantly affects trust, and GEN only significantly affects perceived risk.

3.3. Effective indicators of perceived risk and trust

Participants' braking signal, pupil diameter, and ECG were recorded during the experiment, along with the continuous rating of perceived risk (Fig. 7). This section investigated whether these signals reflect perceived risk and trust effectively.

3.3.1. Continuous ratings of perceived risk

To examine the consistency between continuous and verbal ratings, we compared the participant-averaged peak continuous risk to the corresponding verbal ratings (Fig. 6). The two measures have a strong linear relationship, as indicated by their correlation ($r = 0.983$, $p < 0.001$). Hence, we conclude that the continuous ratings accurately reflect participants' perceived risk.

3.3.2. Braking behaviour and perceived risk and trust

Braking is a signal potentially reflecting higher perceived risk and lower trust levels. This was confirmed by adding the maximum braking pedal position (max_B) to the models in Equation (1) and (2), resulting in model 10 (see Table B. 1 and Table B. 2 in

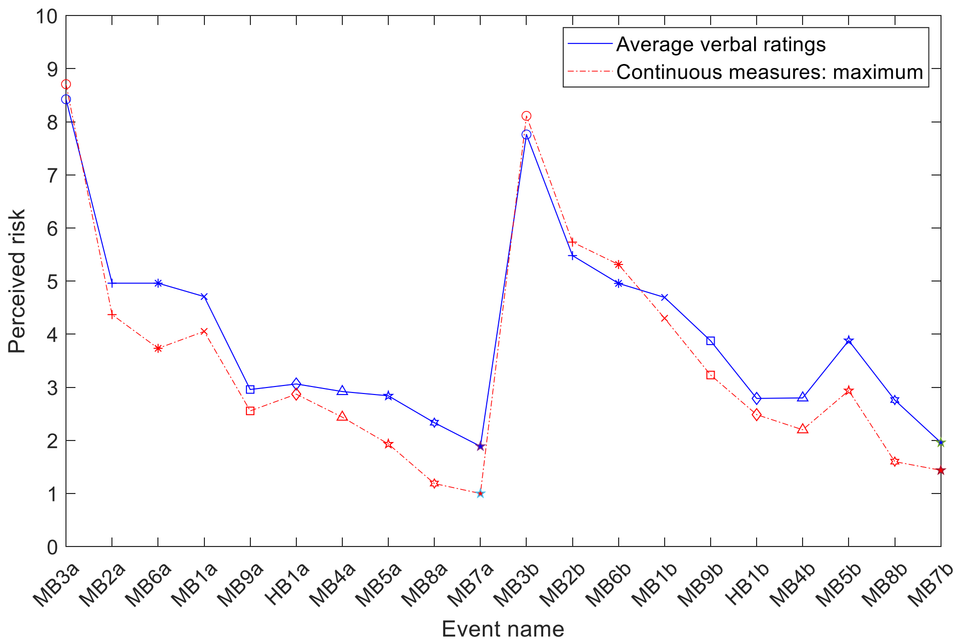

Fig. 6. Participant-averaged perceived risk ratings within events for the two rating methods.

Appendix B) for perceived risk and trust. Maximum braking was a significant predictor of perceived risk ($p = 0.012$), but the relation with trust was not significant ($p = 0.070$) within single events.

7 of 25 participants braked in at least one event. Therefore, the participants were divided into non-braking and braking groups (see Appendix D Braking behaviour). The braking group reported a higher perceived risk ($p = 0.021$) and lower trust ($p = 0.000$) in most of the event types supported by *t*-test.

3.3.3. Pupil dilation and perceived risk

We evaluated if pupil dilation can indicate perceived risk. Pupil dilation was expressed as the difference of the maximum and minimum pupil diameter from 20 s before till 10 s after reaching a minimum gap to the lead vehicle, which spans the first moment when the on-ramp became visible until the time the participant was asked to provide subjective ratings (see Fig. 7).

Firstly, we explore whether the pupil significantly dilated in different event conditions. Kruskal-Wallis tests showed significant variations of participants' pupil diameter within most of the event conditions ($p = 0.000$), indicating pupil dilation significantly changed within an event (see Appendix D Pupil dilation).

To evaluate whether event criticality had a within-subjects effect on pupil dilation, we performed repeated measures ANOVA among all 20 event types. The difference between pupil dilation in different event types is statistically significant in the second exposure of the events ($F\,(5.881, 123.502) = 2.783, p = 0.015$), but insignificant in the first encounters ($F\,(4.470, 93.865) = 1.014, p =$

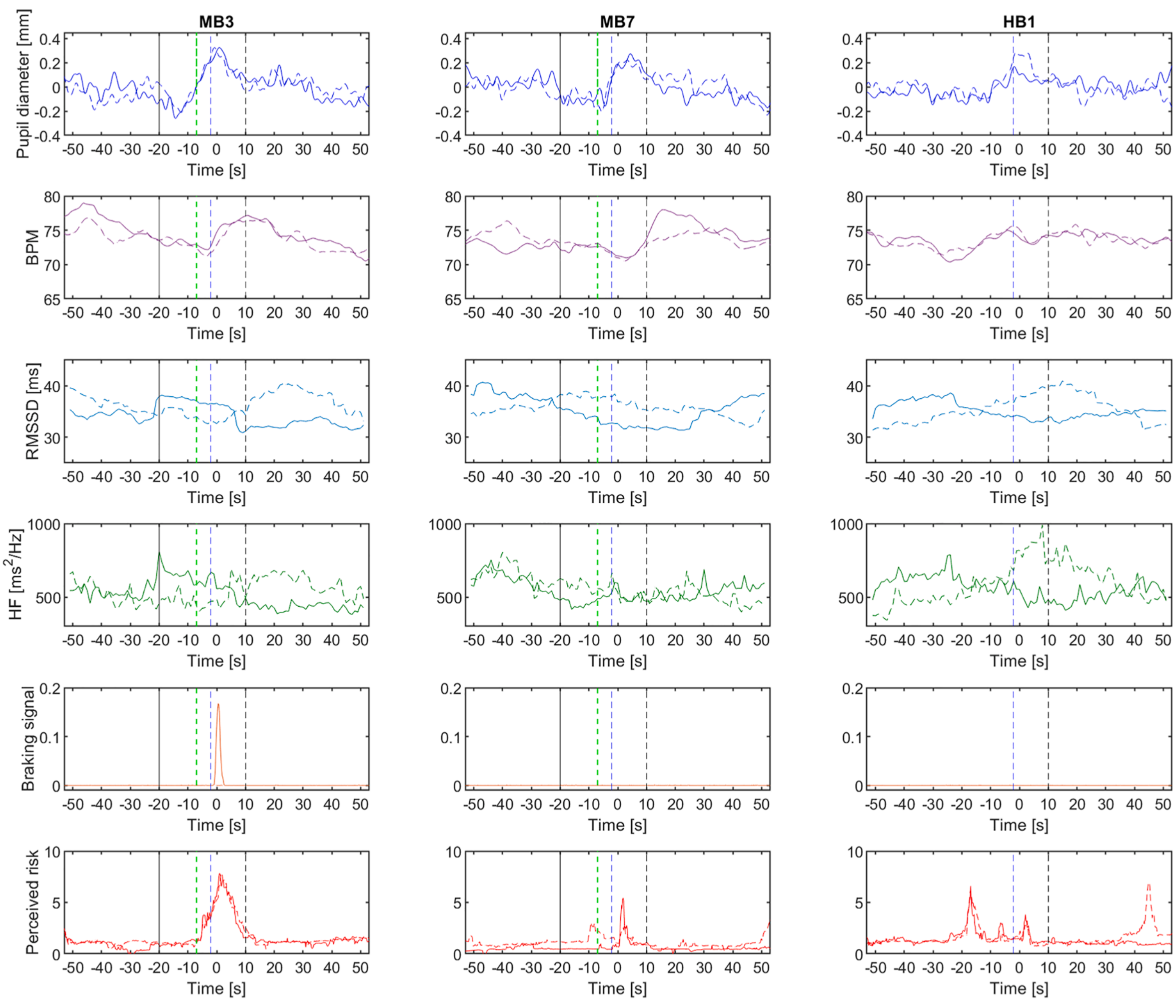

Fig. 7. Participant-averaged signals in the most critical merge MB3 (left), the least critical merge MB7 (mid), and in hard braking (right). The time scale is t = -55 s to t = 55 s, where t = 0 means the smallest gap to the lead vehicle. Pupil diameter is relative to the participant's overall pupil average (all events combined). BPM indicates heat rate; RMSSD represents root mean square of successive inter-beat interval differences; HF indicates the power in the High-frequency band (HF; 0.15–0.40 Hz); The solid curve and the dashed curve represent the first and second exposure. Time is zero for the timing of minimum gap to the merging or leading vehicle. Black solid line: timing when the on-ramp became visible. Green dashed line: timing when the merging vehicles became visible. Blue dashed line: timing when the merging vehicle started to brake. Black dashed line: timing when the participants gave verbal ratings. Results before t = 0 include recovery from preceding events with some variation due to randomization. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

0.409). Repeated measures correlation analysis in 397 events shows no significant correlation between within-subject pupil dilation and the maximum continuous risk (Correlation $r = 0.03$, $p = 0.530$). A strong correlation is found between the participant-averaged pupil diameter signal and the participant-averaged continuous perceived risk signal across all participants (22 participants available in 18 out of the 20 event types, especially in the most critical event MB3 (see Table D. 1 in Appendix D Pupil dilation).

Therefore, we conclude based on the participant-averaged signals that pupil dilation highly correlates with perceived risk if the events are sufficiently risky. The merging and braking events affected pupil dilation, but the correlation between the maximum continuous perceived risk and pupil dilation across all events was not statistically significant.

3.3.4. ECG and perceived risk

To evaluate whether event criticality had a within-subjects effect on heart measures, we performed repeated measures ANOVA as well as a repeated measures correlation among all event types. Heart rate and variability metrics were calculated over the same time period as the pupil dilation in Section 3.3.3. The ANOVA results show that a difference between event types was only statistically observed for IBI with a marginal significance of $p = 0.053$ among first events and $p = 0.002$ among second events. In the most critical event MB3, IBI tended to be smaller compared to less critical events (see Table D. 2 in Appendix D ECG).

Repeated measures correlations were performed between the three heart metrics and two safety metrics: the minimum time headway and maximum continuous risk. A strong correlation between the safety metrics was observed but they do not correlate with the within-subject heart metrics (see Table D. 3 in Appendix D ECG)

In terms of the participant-averaged BPM, RMSSD and HF, heart rate increases in the period following an event, and this pattern is consistent across events and within repetitions of the same event. No such pattern was observed for RMSSD and HF in the participant-averaged signals. The participant-Averaged BPM increase did not correlate significantly with participant-averaged perceived risk ($r = 0.174$, $n = 20$, $p = 0.464$) or minimum time headway ($r = -0.378$, $n = 20$, $p = 0.100$), which indicates no consistent relation with event criticality (see Fig. D. 3 in Appendix B).

Therefore, we conclude that the merging and brake events can increase heart rate, but no significant relation was found between these heart measures and perceived or objective risk. Heart rate variability metrics were not significantly affected.

4. Discussion

This study was conducted to model perceived risk and trust in SAE Level 2 driving automation (ACC + LC) reacting to merging and hard braking vehicles based on a driving simulator experiment.

4.1. Factors influencing perceived risk and trust

The regression models of perceived risk and trust (Equation (1) and (2)) demonstrate that smaller minimum gap, minimum TTC, and stronger braking intensity lead to higher perceived risk and lower trust. Other classic surrogate metrics of safety (SMoS), including the initial merging gap and THW have similar performance predicting perceived risk and trust (Table B. 1 and Table B. 2 in Appendix B).

Our study shows that these well-known SMoS are good predictors for perceived risk. The results are in line with Ma et al. (2018), who used TTC and THW in a regression model of driver's subjective risk in critical cut-in scenarios in manual naturalistic driving. Kondoh et al. (2008) captured human risk perception during car-following as the summation of the time headway inverse and the time to collision. Lu et al. (2012) also found a strong relationship between TTC, THW and perceived risk in car-following. It should be noted that non-linear transformation of the SMoS was required to obtain accurate models predicting perceived risk and trust. These transformations may express human risk perception related to the visual perception of relative motion (Lee, 1976). Braking intensity of the leading vehicle affects both perceived risk and trust as an independent factor which is only mildly correlated to other factors related to vehicle motion ($|r| < 0.275$ for Cluster 2 in Table 3). We are not aware of other studies indicating the relevance of braking intensity in perceived risk and trust, and recommend further exploration of this factor in future studies.

For trust, our study shows that the well-known SMoS are also predictive. Specifically, participants have a lower post-event trust if the minimum TTC is smaller and the braking is stronger in the previous event. This aligns with existing studies on trust modelling and trust calibration. According to Hoff and Bashir (2015) and Lee and See (2004), trust is established in a dynamic process using new information (e.g., event criticality, system performance, etc.). Kraus, Scholz, Stiegemeier, et al. (2020) demonstrate that drivers perceive and interpret the system behaviour and then update the dynamic learned trust based on the initial dynamic trust. Kaplan, Kessler, and Hancock (2021) established a trust model where the environmental or contextual factors are directly used to calculate trust. In our models, the safety metrics and kinematic vehicle state represent the criticality of the previous event, which updates trust based on the initial trust.

The impacts of personal characteristics on perceived risk and trust were quantified by the regression models in Table 4 and Table 5. We found increased trust in automation with more driving experience similar to Jin et al. (2020), Gold et al. (2015) and Kaplan, Kessler, Brill, et al. (2021). For perceived risk, we found that driving-experienced participants perceived lower risk when using automation compared to inexperienced drivers. Apart from a higher trust in the system, this may also be a consequence of trust in their own ability to intervene, as illustrated by He and Donmez (2019), where experienced drivers showed less attention relaxation with SAE Level 2 automation while performing a secondary-task. Similarly, Borowsky and Oron-Gilad (2013) and Ping et al. (2018) show that experienced drivers have stronger hazard awareness when watching movies of real-world driving situations.

Gender effects were found in several studies, where males tend to trust the automation system more than females (Feldhütter et al., 2016; Kaplan, Kessler, Brill, et al., 2021), and females perceive more risk (Hulse et al. (2018). We found females to experience a higher risk but found no significant effect on trust, possibly due to limited sample size and imperfect gender balance.

4.2. The relation between perceived risk and trust

Our results support that when people perceive lower risk, they trust automation more. Models 4 in Table 4 and Table 5 show considerable overlap between trust and perceived risk determinants, confirmed through cross-validation with models 9 in Table B. 1 and Table B. 2 in Appendix B. The two constructs shared a strong negative correlation, which agrees with internet surveys (e.g. Choi and Ji, 2015; Zoellick et al., 2019). Consequently, we observed that predicting variables of either construct are interchangeable between regression models. However, there are grounds to believe that the two constructs are partially independent (see the discussion in 4.4).

4.3. Measures of perceived risk and trust

The regression models are based on post-event verbal ratings of perceived risk and trust. Continuous perceived risk was highly correlated to the post-event verbal rating, indicating the effectiveness of the continuous measurement of perceived risk. However, the continuous measurement increases the drivers' workload, and participants occasionally forget to press the sensor. Taking into account these limitations we do consider our continuous risk measure to be a valuable reference to study and model perceived risk as a function of time. However, in a recent on-road experiment such a measurement was less effective with many missed events (Stapel, Gentner, & Happee, 2022).

Significant effects were obtained for physiological measurements of pupil diameter and ECG metrics. Effects of events on pupil dilation were significant at the group level. Pupil dilation varied with perceived risk, in particular for the most critical events. ECG metrics showed significant effects of events on IBI, but showed no significant correlations with event criticality. Hence, pupil dilation may indicate the amount of perceived risk while IBI is at best indicative of the presence of perceived risk. However, they lack accuracy in quantifying perceived risk which will be an even larger drawback in on-road studies.

We found that driver intervention (braking) relates to perceived risk and trust at the event level, where the 7 participants braking in at least one event reported higher risk and lower trust averaged over all events. Such interventions demonstrate active monitoring by drivers and the somewhat lower trust levels in braking drivers can reflect well-calibrated trust levels. Our findings regarding braking behaviour are in line with Tenhundfeld et al. (2020) and Lee et al. (2021), who found that if participants brake more frequently, they perceive more risk and trust the system less in a Tesla automated parking test and an intersection crossing experiment in a driving simulator. Hence, braking behaviour is a relevant indicator of the presence of perceived risk and distrust but only 7 out of 25 participants braked. Braking can be easily used for research since it will not influence behaviour and requires no additional instrumentation.

4.4. Limitations and future work

Our regression models of perceived risk and trust are based on Level 2 driving automation (ACC + LC) in limited samples (428 events). The models use surrogate metrics of safety related to longitudinal interaction (1-D) to predict perceived risk and trust for both merging with hard braking (MB) and hard braking without merging (HB). However, in the real world, perceived risk comes not only from longitudinal but also from lateral conflicts. More advanced surrogate metrics of safety already consider lateral motion and multiple risk sources (Afghari et al., 2018; Kolekar et al., 2020; Mullakkal-Babu et al., 2020). Further experiments should consider lateral interaction to extend the current models to 2-D. The models predict perceived risk and trust per event, and need a transformation for prediction as continuous function of time suitable for real time control.

The regression models of perceived risk and trust combine existing knowledge with human response data. The selection of potential influencers of perceived risk and trust, and the clusters in Table 3 are knowledge-based, while multicollinearity checks and regression are data-driven. Full data-driven methods, like machine learning, can dig out more valuable information from data beyond human experience, and may be promising in future modelling of perceived risk and trust in particular when large datasets become available from multiple experiments and on-road observations.

A close correlation was observed between perceived risk and trust. However, we still cannot answer the causality question '*which one is the determinant of the other*'. The interaction between perceived risk and trust can be studied further using longitudinal data monitoring the process of trust calibration in relation to event criticality and performance of the driving automation. We found that participants trust the automation more in the second exposure to events, which is shown as REP in trust regression model (Equation (2)). This is related to trust calibration and is also supported by Kraus, Scholz, Stiegemeier, et al. (2020), who found that trust increases over the course of system interaction if an automated system works without malfunctions. During our experiment, no malfunctions or accidents were simulated. Therefore, we can only conclude that participants trust the system more with more automated system interaction, but our results cannot support accident-related trust calibration. The current results only support short-term learned trust calibration. Self-reported trust can still change in several weeks, even in several months, according to Walker et al. (2018), where driver's trust toward SAE level 2 cars still changed 2 weeks after the automation experience. Note that the trust here usually refers to dynamic learned trust calibration, which is steadily updated in a dynamic calibration feedback loop according to Hoff and Bashir (2015), Kraus, Scholz, Stiegemeier, et al. (2020) and Lee and See (2004).

It shall also be pointed out that the applied verbal rating procedure captured perceived risk and trust with only two questions for each event. This simplification was needed to capture ratings after each event but can be complemented by more complex ratings in future studies.

The models of perceived risk and trust can be used to calibrate SMoS so that human factors can be included in risk prediction and assessment. These subjectively calibrated SMoS can be subsequently used as cost function, constraint or reference in AV path planning, decision making and controller design (Hu and Wang, 2021). This has the potential to make users feel safe and trust AV, enhancing user acceptance of driving automation.

5. Conclusions

This study investigated perceived risk and trust in Level 2 driving automation (ACC + LC) in motorway driving with a simulator experiment. We developed regression models that accurately predict perceived risk and trust in specific events and the models reveal that neighbouring road users' behaviours (relative motion) significantly influence occupants' perceived risk and trust. No difference was found in perceived risk and trust between merging with hard braking and hard braking without merging. Our models show that experienced drivers are less sensitive to risk and trust the automation more, while female participants perceive more risk than males. The findings confirm that perceived risk and trust are highly correlated. The proposed models indicate that trust and perceived risk shared the predictors of minimum gap, minimum TTC, years of driving, and braking intensity but differ in using event repetition in the trust model and gender in the perceived risk model. Additionally, the results show that people who perceive lower risk trust the automation more. Regarding the indicators of perceived risk and trust, continuous ratings of perceived risk and braking behaviour can effectively indicate perceived risk or trust. Pupil dilation can reflect perceived risk if the event is sufficiently risky. The merging and braking events increased heart rate, but there was no quantified relation between heart rate increase (variability) and perceived risk.

Future research will focus on extending the perceived risk and trust models towards more complex interactions and applying the models in designing control strategies and human-machine interfaces leading to desirable levels of perceived risk and trust.

CRediT authorship contribution statement

Xiaolin He: Conceptualization, Data curation, Formal analysis, Methodology, Validation, Visualization, Writing – original draft. **Jork Stapel:** Conceptualization, Methodology, Data curation, Formal analysis, Validation, Visualization, Writing – review & editing. **Meng Wang:** Supervision, Conceptualization, Methodology, Writing – review & editing. **Riender Happee:** Supervision, Conceptualization, Methodology, Writing – review & editing.

Declaration of Competing Interest

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

Acknowledgement

This research is supported by Toyota Motor Europe NV/SA and by the SHAPE-IT project funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement 860410.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trf.2022.02.016.

References

Afghari, A. P., Washington, S., Haque, M. M., & Li, Z. (2018). A comprehensive joint econometric model of motor vehicle crashes arising from multiple sources of risk. *Analytic Methods in Accident Research, 18*, 1–14. https://doi.org/10.1016/j.amar.2018.03.002

Ajenaghughrure, I. B., Sousa, S. C. D. C., & Lamas, D. (2021). In *Psychophysiological modelling of trust in technology: Comparative analysis of algorithm ensemble methods* (pp. 000161–000168). IEEE. https://doi.org/10.5220/0010237701610173.

Astin, A. D. (1999). *Finger force capability: Measurement and prediction using anthropometric and myoelectric measures*. USA: Virginia Polytechnic Institute and State University. Ph.D. Thesis.

Ayoub, J., Yang, X. J., & Zhou, F. (2021). Modeling dispositional and initial learned trust in automated vehicles with predictability and explainability. *Transportation Research Part F: Traffic Psychology and Behaviour, 77*, 102–116. https://doi.org/10.1016/j.trf.2020.12.015

Baek, H. J., Cho, C. H., Cho, J., & Woo, J. M. (2015). Reliability of ultra-short-term analysis as a surrogate of standard 5-min analysis of heart rate variability. *Telemedicine and e-Health, 21*(5), 404–414. https://doi.org/10.1089/tmj.2014.0104

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology, 8*, 456. https://doi.org/10.3389/fpsyg.2017.00456

Beggiato, M., Hartwich, F., & Krems, J. (2019). Physiological correlates of discomfort in automated driving. *TransportationResearch Part F: Traffic Psychology and Behaviour, 66*, 445–458. https://doi.org/10.1016/j.trf.2019.09.018

Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour, 35*, 75–84. https://doi.org/10.1016/j.trf.2015.10.005

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods, 190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Borowsky, A., & Oron-Gilad, T. (2013). Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. *Accident Analysis & Prevention, 59*, 548–565. https://doi.org/10.1016/j.aap.2013.07.008

Choi, J. K., & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction, 31*(10), 692–702. https://doi.org/10.1080/10447318.2015.1070549

Cleij, D., Venrooij, J., Pretto, P., Pool, D. M., Mulder, M., & Bülthoff, H. H. (2018). Continuous subjective rating of perceived motion incongruence during driving simulation. *IEEE Transactions on Human-Machine Systems, 48*(1), 17–29. https://doi.org/10.1109/THMS.2017.2717884

Cox, J. A., Beanland, V., & Filtness, A. J. (2017). Risk and safety perception on urban and rural roads: Effects of environmental features, driver age and risk sensitivity. *Traffic Injury Prevention, 18*(7), 703–710. https://doi.org/10.1080/15389588.2017.1296956

Dong, X., DiScenna, M., & Guerra, E. (2019). Transit user perceptions of driverless buses. *Transportation, 46*(1), 35–50. https://doi.org/10.1007/s11116-017-9786-y

Dreger, F. A., de Winter, J. C., & Happee, R. (2020). How do drivers merge heavy goods vehicles onto freeways? A semi-structured interview unveiling needs for communication and support. *Cognition, Technology & Work, 22*(4), 825–842. https://doi.org/10.1007/s10111-019-00601-3

ERTRAC (2017). Automated Driving Roadmap.

Feldhütter, A., Gold, C., Hüger, A., & Bengler, K. (2016). Trust in automation as a matter of media influence and experi-ence of automated vehicles. In Proceedings of the human factors and ergonomics society annual meeting, Los Angeles, Vol. 60, No. 1, pp. 2024-2028. https://doi.org/10.1177/1541931213601460.

Fuller, R. (2011). Driver control theory: From task difficulty homeostasis to risk allostasis. In Handbook of traffic psychology, pp. 13-26. https://doi.org/10.1016/B978-0-12-381984-0.10002-5.

Garidis, K., Ulbricht, L., Rossmann, A., & Schmäh, M. (2020). Toward a user acceptance model of autonomous driving. In Proceedings of the 53rd Hawaii international conference on system sciences. Geller, E.S., 2016. The psychology of safety handbook. CRC press.

Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in automation–before and after the experience of take-over scenarios in a highly automated vehicle. *Procedia Manufacturing, 3*, 3025–3032. https://doi.org/10.1016/j.promfg.2015.07.847

Griffin, W., Haworth, N., & Twisk, D. (2020). Patterns in perceived crash risk among male and female drivers with and without substantial cycling experience. *Transportation Research Part F: Traffic Psychology and Behaviour, 69*, 1–12. https://doi.org/10.1016/j.trf.2019.12.013

Hancock, P. A., Kessler, T. T., Kaplan, A. D., Brill, J. C., & Szalma, J. L. (2020). Evolving trust in robots: Specification through sequential and comparative meta-analyses. *Human Factors*. https://doi.org/10.1177/0018720820922080

He, D., & Donmez, B. (2019). Influence of driving experience on distraction engagement in automated vehicles. *Transportation Research Record, 2673*(9), 142–151. https://doi.org/10.1177/0361198119843476

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Hu, C., & Wang, J. (2021). Trust-Based and Individualizable Adaptive Cruise Control Using Control Barrier Function Approach With Prescribed Performance. *IEEE Transactions on Intelligent Transportation Systems*. https://doi.org/10.1109/TITS.2021.3066154

Huang, L., Xia, Q., Xie, F., Xiu, H. L., & Shu, H. (2018). In *Study on the test scenarios of level 2 automated vehicles* (pp. 49–54). IEEE. https://doi.org/10.1109/IVS.2018.8500600.

Hulse, L. M., Xie, H., & Galea, E. R. (2018). Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety Science, 102*, 1–13. https://doi.org/10.1016/j.ssci.2017.10.001

ISO (2018a). Road vehicles – Functional safety – Part 1: Vocabulary.

ISO (2018b). Intelligent transport systems — Adaptive cruise control systems — Performance requirements and test.

Jin, M., Lu, G., Chen, F., & Shi, X. (2020). In *How Driving Experience Affect Trust in Automation from Level 3 Automated Vehicles? An Experimental Analysis* (pp. 1–6). IEEE. https://doi.org/10.1109/ITSC45102.2020.9294247.

Kaplan, A. D., Kessler, T. T., Brill, J. C., & Hancock, P. A. (2021). Trust in Artificial Intelligence: Meta-Analytic Findings. *Human Factors*. https://doi.org/10.1177/00187208211013988

Kaplan, A. D., Kessler, T. T., & Hancock, P. A. (2020). How Trust is Defined and its use in Human-Human and Human-Machine Interaction. *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Los Angeles, 64*(1), 1150–1154. https://doi.org/10.1177/1071181320641275

Kaplan, A. D., Kessler, T. T., Sanders, T. L., Cruit, J., Brill, J. C., & Hancock, P. A. (2021). A time to trust: Trust as a function of time in human-robot interaction. *Trust in Human-Robot Interaction*. Elsevier Inc.. https://doi.org/10.1016/b978-0-12-819472-0.00006-x.

Khusro, Y. R., Zheng, Y., Grottoli, M., & Shyrokau, B. (2020). MPC-Based Motion-Cueing Algorithm for a 6-DOF Driving Simulator with Actuator Constraints. *Vehicles, 2*(4), 625–647. https://doi.org/10.3390/vehicles2040036

Kiefer, R. J., LeBlanc, D. J., & Flannagan, C. A. (2005). Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments. *Accident Analysis & Prevention, 37*(2), 295–303. https://doi.org/10.1016/j.aap.2004.09.003

Kolekar, S., de Winter, J., & Abbink, D. (2020). Which parts of the road guide obstacle avoidance? Quantifying the driver's risk field. *Applied ergonomics, 89*, Article 103196. https://doi.org/10.1016/j.apergo.2020.103196

Kondoh, T., Yamamura, T., Kitazaki, S., Kuge, N., & Boer, E. (2008). Identification of Visual Cues and Quantification of Drivers' Perception of Proximity Risk to the Lead Vehicle in Car-Following Situations. *Journal of Mechanical Systems for Transportation and Logistics, 1*(2), 170–180. https://doi.org/10.1299/jmtl.1.170

Kraus, J., Scholz, D., & Baumann, M. (2020). What's Driving Me? Exploration and Validation of a Hierarchical Personality Model for Trust in Automated Driving. *Human Factors.*. https://doi.org/10.1177/0018720820922653

Kraus, J., Scholz, D., Messner, E. M., Messner, M., & Baumann, M. (2020). Scared to trust?–predicting trust in highly automated driving by depressiveness, negative self-evaluations and state anxiety. *Frontiers in Psychology, 10*, 2917. https://doi.org/10.3389/fpsyg.2019.02917

Kraus, J., Scholz, D., Stiegemeier, D., & Baumann, M. (2020). The more you know: Trust dynamics and calibration in highly automated driving and the effects of take-overs, system malfunction, and system transparency. *Human Factors, 62*(5), 718–736. https://doi.org/10.1177/0018720819853686

Kret, M. E., & Sjak-Shie, E. E. (2019). Preprocessing pupil size data: Guidelines and code. *Behavior Research Methods, 51*(3), 1336–1342. https://doi.org/10.3758/s13428-018-1075-y

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception, 5*(4), 437–459.

Lee, J. D., Liu, S. Y., Domeyer, J., & DinparastDjadid, A. (2021). Assessing drivers' trust of automated vehicle driving styles with a two-part mixed model of intervention tendency and magnitude. *Human Factors, 63*(2), 197–209. https://doi.org/10.1177/0018720819880363

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50.30392

Li, M., Holthausen, B. E., Stuck, R. E., & Walker, B. N. (2019). No risk no trust: Investigating perceived risk in highly automated driving. In *In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 177–185). https://doi.org/10.1145/3342197.3344525

Lu, G., Cheng, B., Lin, Q., & Wang, Y. (2012). Quantitative indicator of homeostatic risk perception in car following. *Safety Science, 50*(9), 1898–1905. https://doi.org/10.1016/j.ssci.2012.05.007

Ma, X., Feng, Z., Zhu, X., & Ma, Z. (2018). Driver Risk Perception Model under Critical Cut-In Scenarios (No. 2018-01-1626). SAE Technical Paper. https://doi.org/10.4271/2018-01-1626.

Marsh, S., & Dibben, M. R. (2003). The role of trust in information science and technology. *Annual Review of Information Science and Technology (ARIST), 37*, 465–498.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20*(3), 709–734.

Meyer-Waarden, L., & Cloarec, J. (2022). "Baby, you can drive my car": Psychological antecedents that drive consumers' adoption of AI-powered autonomous vehicles. Technovation, 109(April 2021), 102348. https://doi.org/10.1016/j.technovation.2021.102348.

Morris, D. M., Erno, J. M., & Pilcher, J. J. (2017). Electrodermal response and automation trust during simulated self-driving car use. *In Proceedings of the human factors and ergonomics society annual meeting, Los Angeles, 61*(1), 1759–1762. https://doi.org/10.1177/1541931213601921

Mullakkal-Babu, F. A., Wang, M., He, X., van Arem, B., & Happee, R. (2020). Probabilistic field approach for motorway driving risk assessment. *Transportation Research Part C: Emerging Technologies, 118*, Article 102716. https://doi.org/10.1016/j.trc.2020.102716

Mullakkal-Babu, F. A., Wang, M., Van Arem, B., & Happee, R. (2016). Design and analysis of full range adaptive cruise control with integrated collision a voidance strategy. In 2016 IEEE 19th International conference on intelligent transportation systems (ITSC), pp. 308-315.https://doi.org/10.1109/ITSC.2016.7795572.

Nordhoff, S., Stapel, J., He, X., Gentner, A., & Happee, R. (2021). Perceived safety and trust in SAE Level 2 partially automated cars : Results from an online questionnaire. *PLoS ONE, 16*(12). https://doi.org/10.1371/journal.pone.0260953

Nordhoff, S., Malmsten, V., van Arem, B., Liu, P., & Happee, R. (2021). A structural equation modeling approach for the acceptance of driverless automated shuttles based on constructs from the Unified Theory of Acceptance and Use of Technology and the Diffusion of Innovation Theory. *Transportation Research Part F: Traffic Psychology and Behaviour, 78*, 58–73. https://doi.org/10.1016/j.trf.2021.01.001

Noy, I. Y., Shinar, D., & Horrey, W. J. (2018). Automated driving: Safety blind spots. *Safety Science, 102*, 68–78. https://doi.org/10.1016/j.ssci.2017.07.018

Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). Predicting information technology usage in the car: Towards a car technology acceptance model. In *In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 51–58). https://doi.org/10.1145/2390256.2390264

Perello-March, J., Burns, C., Elliott, M., & Birrell, S. (2019). Integrating trust in automation into driver state monitoring systems. *International Conference on Human Interaction and Emerging Technologies, 344–349*. https://doi.org/10.1007/978-3-030-25629-6

Ping, P., Sheng, Y., Qin, W., Miyajima, C., & Takeda, K. (2018). Modeling driver risk perception on city roads using deep learning. *IEEE Access, 6*, 68850–68866. https://doi.org/10.1109/ACCESS.2018.2879887

Pyrialakou, V. D., Gkartzonikas, C., Gatlin, J. D., & Gkritza, K. (2020). Perceptions of safety on a shared road: Driving, cycling, or walking near an autonomous vehicle. *Journal of Safety Research, 72*, 249–258. https://doi.org/10.1016/j.jsr.2019.12.017

Rhodes, N., & Pivik, K. (2011). Age and gender differences in risky driving: The roles of positive affect and risk perception. *Accident Analysis & Prevention, 43*(3), 923–931. https://doi.org/10.1016/j.aap.2010.11.015

Rossner, P., & Bullinger, A. C. (2019). How do you want to be driven? investigation of different highly-automated driving styles on a highway scenario. In *International Conference on Applied Human Factors and Ergonomics* (pp. 36–43). Cham: Springer.

SAE. (2021). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*. Warrendale, PA, USA: SAE International.

Saffarian, M., Happee, R., & Winter, J. D. (2012). Why do drivers maintain short headways in fog? A driving-simulator study evaluating feeling of risk and lateral control during automated and manual car following. *Ergonomics, 55*(9), 971–985. https://doi.org/10.1080/00140139.2012.691993

Schaefer, K. (2013). The Perception and Measurement of Human-Robot Trust. Ph.D. Thesis, University of Central Florida, USA. https://doi.org/CFE0004931.

Sedghamiz, H. (2018). BioSigKit: A matlab toolbox and interface for analysis of biosignals. *Journal of Open Source Software, 3*(30), 671. https://doi.org/10.21105/joss.00671.

Sedghamiz, H., & Santonocito, D. (2015). In *Unsupervised detection and classification of motor unit action potentials in intramuscular electromyography signals* (pp. 1–6). IEEE. https://doi.org/10.1109/EHB.2015.7391510.

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health, 258*. https://doi.org/10.3389/fpubh.2017.00258

Sharma, A., Zheng, Z., Kim, J., Bhaskar, A., & Haque, M. M. (2020). Is an informed driver a better decision maker? A grouped random parameters with heterogeneity-in-means approach to investigate the impact of the connected environment on driving behaviour in safety-critical situations. *Analytic Methods in Accident Research, 27*, Article 100127. https://doi.org/10.1016/j.amar.2020.100127

Stapel, J., Mullakkal-Babu, F. A., & Happee, R. (2019). Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 60*, 590–605. https://doi.org/10.1016/j.trf.2018.11.006

Stapel, J., Gentner, A., & Happee, R. (2022). On-road trust and perceived safety in level 2 automation. Manuscript in preparation.

Tang, R., Kim, J. H., Parker, R., & Jeong, Y. J. (2018). Indicating severity of vehicle accidents using pupil diameter in a driving simulator environment. In *International Conference on Digital Human Modeling and Applications in Health* (pp. 647–656). Ergonomics and Risk Management: Safety.

Taylor, D. H. (1964). Drivers' Galvanic Skin Response and The Risk of Accident. *Ergonomics, 7*(4), 439–451. https://doi.org/10.1080/00140136408930761

Tenhundfeld, N. L., de Visser, E. J., Ries, A. J., Finomore, V. S., & Tossell, C. C. (2020). Trust and distrust of automated parking in a Tesla Model X. *Human Factors, 62*(2), 194–210. https://doi.org/10.1177/0018720819865412

Varotto, S. F., Farah, H., Toledo, T., van Arem, B., & Hoogendoorn, S. P. (2018). Modelling decisions of control transitions and target speed regulations in full-range Adaptive Cruise Control based on Risk Allostasis Theory. *Transportation Research Part B: Methodological, 117*, 318–341. https://doi.org/10.1016/j.trb.2018.09.007

Walker, F., Boelhouwer, A., Alkim, T., Verwey, W. B., & Martens, M. H. (2018). Changes in trust after driving level 2 automated cars. *Journal of Advanced Transportation, 2018*. https://doi.org/10.1155/2018/1045186

Wang, C., Xie, Y., Huang, H., & Liu, P. (2021). A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling. *Accident Analysis & Prevention, 157*, Article 106157. https://doi.org/10.1016/j.aap.2021.106157

Xiong, X., Wang, M., Cai, Y., Chen, L., Farah, H., & Hagenzieker, M. (2019). A forward collision avoidance algorithm based on driver braking behavior. *Accident Analysis & Prevention, 129*, 30–43. https://doi.org/10.1016/j.aap.2019.05.004

Xu, Z., Zhang, K., Min, H., Wang, Z., Zhao, X., & Liu, P. (2018). What drives people to accept automated vehicles? Findings from a field experiment. *Transportation Research Part C: Emerging Technologies, 95*, 320–334. https://doi.org/10.1016/j.trc.2018.07.024

Yagoda, R. E., & Gillan, D. J. (2012). You want me to trust a ROBOT? The development of a human–robot interaction trust scale. *International Journal of Social Robotics, 4*(3), 235–248. https://doi.org/10.1007/s12369-012-0144-0

Zhao, D., Guo, Y., & Jia, Y. J. (2017). TrafficNet: An open naturalistic driving scenario library. In *In 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), IEEE* (pp. 1–8).

Zoellick, J. C., Kuhlmey, A., Schenk, L., Schindel, D., & Blüher, S. (2019). Amused, accepted, and used? Attitudes and emotions towards automated vehicles, their relationships, and predictive value for usage intention. *Transportation Research Part F: Traffic Psychology and Behaviour, 65*, 68–78. https://doi.org/10.1016/j.trf.2019.07.009