



Universiteit
Leiden
The Netherlands

Do engineer perceptions about automated vehicles match user trust? Consequences for design

Walker, F.; Steinke, J.; Martens, M.H.; Verwey, W.B.

Citation

Walker, F., Steinke, J., Martens, M. H., & Verwey, W. B. (2022). Do engineer perceptions about automated vehicles match user trust? Consequences for design. *Transportation Research Interdisciplinary Perspectives*, 8. doi:10.1016/j.trip.2020.100251

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3480056>

Note: To cite this publication please use the final published version (if applicable).



Do engineer perceptions about automated vehicles match user trust? Consequences for design



F. Walker^{a,*}, J. Steinke^a, M.H. Martens^{b,c}, W.B. Verwey^a

^a University of Twente, Drienerlolaan 5, 7522 NB Enschede, The Netherlands

^b Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands

^c TNO, Anna van Buerenplein 1, 2595 DA The Hague, The Netherlands

ARTICLE INFO

Keywords:

Trust
Trust in automation
Trust calibration
Appropriate trust
Automated vehicle
Designer-user mismatch

ABSTRACT

To maximize road safety, driver trust in an automated vehicle should be aligned with the vehicle's technical reliability, avoiding under- and over-estimation of its capabilities. This is known as trust calibration. In the study reported here, we asked how far participant assessments of vehicle capabilities aligned with those of the engineers. This was done by asking the engineers to rate the reliability of the vehicle in a specific set of scenarios. We then carried out a driving simulator study using the same scenarios, and measured participant trust. The results suggest that user trust and engineer perceptions of vehicle reliability are often misaligned, with users sometimes under-trusting and sometimes over-trusting vehicle capabilities. On these bases, we formulated recommendations to mitigate under- and over-trust. Specific recommendations to improve trust calibration include the adoption of a more defensive driving style for first-time users, the visual representation of the objects detected by the automated driving system in its surroundings in the Human Machine Interface, and real-time feedback on the performance of the technology.

1. Introduction

Automated vehicles promise to drastically reduce road accidents, increase travelling comfort, and reduce driver workload (Fagnant and Kockelman, 2015; Kyriakidis et al., 2019; Litman, 2017; Milakis et al., 2017; Payre et al., 2016; Urmsion and Whittaker, 2008). Nevertheless, none of these benefits will be achieved if users do not trust the technology.

According to Lee and See's (2004) widely used definition, trust is "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (Lee and See, 2004, p. 51). "Uncertainty" and "vulnerability" are key elements in this definition: on the one hand trust is always linked to an uncertain outcome, on the other, perceptions of risk play a crucial role in its development (Brower et al., 2000; Hoff and Bashir, 2015; Lee and See, 2004; Nyhan, 2000; Perkins et al., 2010; Shapiro, 1987). This is true not only in relationships between individuals, but also in relationships between humans and automated driving systems.

The literature shows that trust influences the intention to adopt automated vehicles (Choi and Ji, 2015; Ghazizadeh et al., 2012; Parasuraman and Riley, 1997) and that it plays a fundamental role

in determining a positive user-experience (Ekman et al., 2019; Waytz et al., 2014). Trust requires an affective evaluation of the (perceived) characteristics of the automated vehicle, such as its reliability and thus ability to perform certain tasks (Körber, 2018). It follows that perceived *performance* can be seen as a crucial and intrinsic dimension of trust (Lee and See, 2004; Mayer et al., 1995).

Trust does not only predict the use, but also the misuse and disuse of automated systems (Hoff and Bashir, 2015; Lee and See, 2004; Parasuraman and Riley, 1997). Too much trust – over-trust – can cause over-reliance on the automated system, creating the risk that the user will operate the system in ways that were not originally intended by its designers. Insufficient trust – under-trust – may arise from disappointing interactions with the automated technology and may prevent users from taking advantage of the system's full capabilities, or even using the system at all (Carsten and Martens, 2019; Lee and See, 2004; Parasuraman and Riley, 1997; Payre et al., 2016). In addition, even when an automated system behaves perfectly in line with the designers' predictions, users may want it to behave differently or to provide feedback to explain its behaviour. To avoid misuse and disuse of the automated system trust should be calibrated, and therefore become fully aligned with the actual reliability of the vehicle (Khastgir et al.,

* Corresponding author.

E-mail address: f.walker@utwente.nl (F. Walker).

2017; Lee and See, 2004; Muir, 1987; Payre et al., 2016; Walker et al., 2018). We define the latter as the probability that, in a defined environment, the automated system will perform as expected by its designers.

The calibration of users' trust represents an important goal for i-CAVE (Integrated Cooperative Automated Vehicle). i-CAVE is a multidisciplinary Dutch research programme, focused on the development of a fleet of cooperative automated concept vehicles to be operated on the campus of the Eindhoven University of Technology (The Netherlands) (i-CAVE, 2020). The car, a modified Renault Twizy, will transport people and goods, and will operate with Level 4 automation (SAE, 2018). Therefore, the vehicle will be able to cope with any environment within a specified Operational Design Domain (ODD) (SAE, 2018). This means that users may at times still need to take over, when the vehicle reaches the ODD limits.

As stated by Lee and See (2004), under- and over-trust may be mitigated by designing for "appropriate" rather than "greater trust", and therefore by acting on the vehicle's behaviour and/or on its Human Machine Interface (HMI). Concerning vehicle dynamics, studies have shown that the automated vehicle's driving style may strongly affect driver trust and comfort (Price et al., 2016; Lee et al., 2016; Ekman et al., 2019). In particular, Price et al. (2016) showed that participants trusted a simulated automated vehicle more when this kept a more centred position on the driving lane. Similarly, Lee et al. (2016) pointed out that when participants perceived the lane positioning of the automated vehicle as "imprecise", this negatively affected their trust towards the system. In a recent study using Wizard-of-Oz techniques, Ekman et al. (2019) showed that participants perceived a defensive driving style as more trust-worthy than an aggressive one, preferring an automated vehicle that avoided heavy accelerations, and behaved in a more smooth and predictable way (Ekman et al., 2019).

Concerning the HMI, studies have shown that presenting real-time visual information about the automated vehicle's performance can lead to better trust calibration (Helldin et al., 2013; Kunze et al., 2019). For example, in a driving simulator study by Helldin et al. (2013), the authors presented information on the reliability of the Highly Automated vehicle's behaviour through seven bars that were presented in-car. Each bar indicated the vehicle's ability to keep driving automatically, with 1 indicating "no ability" and 7 "very high ability" (Helldin et al., 2013). Helldin et al.'s (2013) results showed that participants that were presented with the reliability information trusted the system less and spent less time looking at the road. Yet, when needed, they took back control of the car faster than drivers that did not receive such information. More recently, Kunze et al. (2019) confirmed Helldin et al.'s (2013) findings. In this study, participants who received continuous feedback on the performance of the automated driving system calibrated their trust more easily than those who did not. Specifically, in a low visibility situation, they paid more attention to the road, solved fewer non-driving-related-tasks, and reported lower trust scores (Kunze et al., 2019). Notably, the authors' results showed that trust calibration – and safer human-automation interaction – may require *less* rather than more driver trust (Helldin et al., 2013; Kunze et al., 2019).

Helldin et al. (2013) and Kunze et al. (2019) investigated drivers' reactions to specific driving conditions (i.e., situations of low visibility due to snow or fog). They assumed that the reliability of the automated system was known, but this is often not the case. Particularly for automated vehicles equipped with Level 4 driving functions, information concerning the reliability of the system is not even available yet, since these are primarily still being tested in pilots and demonstration projects. The same holds for i-CAVE. At this stage in the i-CAVE design process, the only reliability data available are the judgments of the vehicles' engineers. While a number of studies have investigated how trust influences the interaction with automated driving systems (e.g., Hergeth et al., 2016; Parasuraman et al., 2008; Payre et al.,

2016; Walker et al., 2019), it remains entirely unclear how poor trust calibration can be detected.

In the present study, we asked the engineering team in i-CAVE to estimate the reliability of the automated Twizy in a number of urban driving scenarios. We then recreated these scenarios in our driving simulator, asked participants to experience them, and compared their trust score in each scenario with the engineers' judgments of the car's reliability in those scenarios. Ideally, if the engineers' evaluation shows that the vehicle can handle each scenario, users should trust it, and vice versa. However, if user trust is not aligned with the engineers' judgments, then under- or over-trust may lead to disuse, discomfort or dangerous interactions with the system.

Our goal was exploratory. First, we assessed whether there is a mismatch between first-time users' trust and engineers' judgements of reliability in different driving situations. Second, we aimed to identify factors responsible for trust calibration (i.e., optimal level of trust calibrated to actual vehicle capabilities). Finally, we derived recommendations for vehicle design – to be implemented before actual on-road testing. Such recommendations, although relevant to the i-CAVE vehicle, are particularly important for the calibration of users' trust towards comparable automated driving systems.

2. Methods

2.1. Engineers and participants

The three i-CAVE engineers who were responsible for developing the controllers of the vehicle and its underlying path planning algorithms, participated. The latter are fundamental for the safe deployment of the vehicle in situations of mixed traffic. The same engineers were also responsible for the vehicle's functional architecture and for the evaluation of its safety systems. The focus of their work was on the development of software systems, architectural models and quality standards ensuring the functional safety of the vehicle (i-CAVE, 2020).

Sixty-two participants, all students or employees of the University of Twente, were recruited as "users". They participated in exchange for money (€6) or study credits. None of the sixty-two participants reported previous experience with automated vehicles, and none of them commonly suffered from motion sickness. They all had a driver's licence, usually driving once or twice per week. Mean driving experience was 3.48 years (SD = 3.08). Participants (thirty-four female, twenty-eight male) were between eighteen and forty-one years of age (M = 21.3, SD = 3.5). The study was approved by the ethics board of the Faculty of Behavioural, Management and Social Sciences at the University of Twente.

2.2. Engineers' evaluation

The engineers were asked through an on-line questionnaire to imagine the fully functional i-CAVE vehicle driving automatically in nine urban scenarios (see Fig. 1 and Appendix A). These were all situations commonly experienced by drivers on urban roads (e.g., entering a roundabout, giving right of way at an intersection, overtaking a parked vehicle). For each scenario, they were asked to use their expertise to estimate how *reliably* the car would behave. Scenarios were displayed from a bird-eye view to make sure that engineers' judgments would be based on all the information in the driving environment, and that they would not be influenced by other factors (e.g., trust, feelings of discomfort) that could arise during a simulated drive.

A brief description of the scenario was provided, and optimal weather and road conditions were assumed. Where appropriate, behaviour of pedestrians and other road users was clearly indicated in each figure. The engineers indicated their response on a five-point Likert scale, with "1" indicating minimum reliability and "5" indicating max-

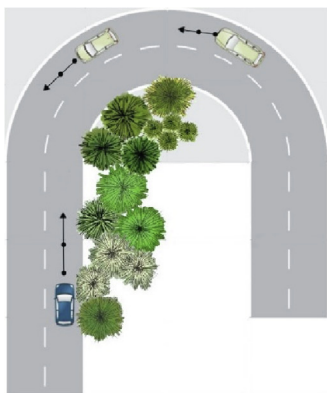


Fig. 1. One of the nine scenarios presented to the engineers (i.e., “Uphill”).

Description given: The automated vehicle is approaching a blind curve (due to bushes on the right-hand side of the road). The road is uphill. Oncoming traffic comes around the curve.

What is the maximum reliability you can achieve in this situation?

(Min) 1 – 2 – 3 – 4 – 5 (Max)

If reliability is below 5, please briefly explain why. What can go wrong in this scenario?

imum reliability. “1” indicated that the automated vehicle could not safely handle the scenario, and that ideally it should never encounter such a situation. “5” indicated that the vehicle could handle the scenario perfectly, and thus that the passenger and the external users in the scenario (e.g., pedestrians, oncoming traffic) had nothing to worry about. When the engineers’ responses were below five, they were asked to briefly explain why through an open question.

2.3. Simulated driving scenarios

The nine scenarios rated by the engineers were recreated in our driving simulator (see Fig. 2). This consists of a skeletal mock-up car positioned in front of a visual screen. The vehicle’s dashboard – an Asus Transformer Book (10.4 × 6.7 in.) – displays speed (in km/h) and a rev counter. When sitting in the driver’s seat, participants experience a 180° field of view (see Fig. 2). Our setup runs with SILAB Version 6.0 software (Wivw GmbH-Silab, 2018) and can be classified as a mid-level driving simulator (Kaptein et al., 1996).

2.4. Procedure and user trust

After collecting participants’ demographic and driving experience information, we measured their general trust in automated vehicles through a modified version of the Empirically Derived (ED) Trust Scale (Jian et al., 2000; Verberne et al., 2012). As in Verberne et al. (2012) and Walker et al. (2019), participants indicated their level of agreement with seven statements (1 = totally disagree; 7 = totally agree). The higher the average score, the higher the trust, and vice versa.

Participants were told that the goal of the study was to assess their trust in the simulated automated vehicle in different driving situations. They were then asked to sit in the driver’s seat and told to supervise

the automated system at all times, although intervention would not be required. A five-minute familiarization phase allowed participants to get used to the simulator. Here, participants experienced a standard automated motorway drive, unrelated to the actual experimental session.

The nine scenarios were then played to the participants in sequence. At the beginning of each scenario, by pressing a button on the steering wheel, participants initiated the vehicle’s automated functionalities. Each scenario started with the simulated automated vehicle driving in an urban environment with no traffic, at a constant speed of 50 km/h. When the automated vehicle encountered the situation of interest, the experiment was paused. Therefore, participants experienced the run-up to the situation, and not the situation itself. Although the paused driving scenario was viewed from a first-person perspective (i.e., driver view), it contained all the visual elements presented also to the engineers.

The experimenter briefly described the situation (using the same description provided to the engineers), and asked each participant “On a scale from one to five, where one is “not at all” and five is “absolutely”, how sure are you that the vehicle can handle this situation?”. Participants wrote down their answer, together with a brief description of the reasons for their rating. Importantly, participants were given no feedback after their responses, which might have affected their subsequent ratings. After the participants’ response, a new scenario was presented. The order of the scenarios was counter-balanced across participants.

After experiencing all nine driving scenarios, participants rated their general trust in the vehicle through the modified ED trust scale and were asked to fill in an exit questionnaire. This was composed of fifteen items, consisting of twelve closed questions (with responses on a five-point ordinal scale) and three open-ended questions (see Appendix B).

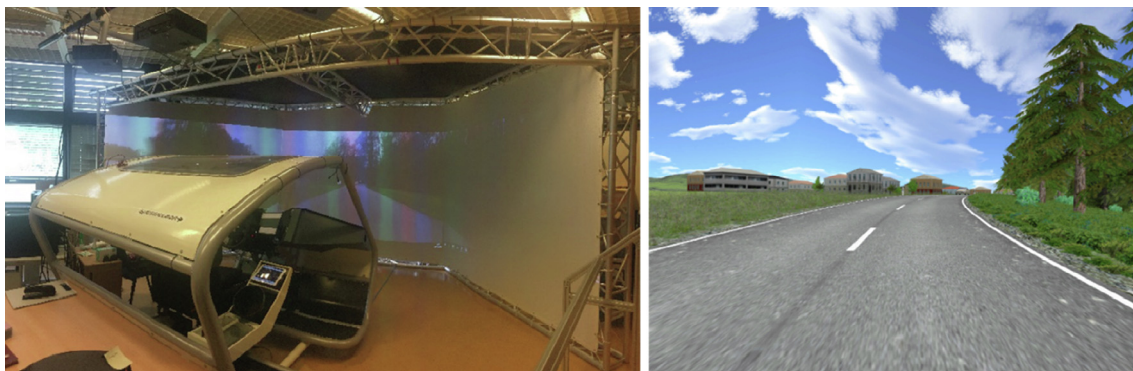


Fig. 2. On the left, the driving simulator of the University of Twente; on the right, the simulated driving environment.

The twelve closed items concerned the vehicle's behaviour (e.g., speed and steering behaviour), and the information provided on its dashboard. The final three open-ended questions allowed participants to indicate what vehicle behaviour did not meet their expectations, if they wanted the vehicle to provide additional information (i.e., feedback), and if there was anything else that they missed in the vehicle – features that could be implemented in the final version of the i-CAVE vehicles.

3. Results

3.1. General trust in the automated vehicle

A related-samples Wilcoxon Signed Ranks test was used to compare participants' general trust in automated vehicles, as measured through the ED trust scale before and after the simulated driving experience. A significant difference was found between pre ($M = 4.13$, $SD = 0.86$) and post ($M = 4.36$, $SD = 1$) trust scores; $Z = -2.35$, $p = .019$. This shows that the simulated on-road experience increased user trust in automation.

3.2. Comparison of reliability and trust scores

With the exception of scenarios Bus and Roundabout, the engineers consistently assigned a high reliability score (4 or 5) to the vehicle. In eight out of nine scenarios (all except scenario Roundabout) the standard deviation of the engineers' responses was less than 1, suggesting agreement among their answers.

To test the alignment between participant and engineer assessments we used a one-sample Wilcoxon Signed Ranks test. This is used to determine whether the median of a sample (participant trust assessments) matches a known value (engineer reliability assessments). A significant difference between the two medians is evidence that participant trust may be misaligned with engineer assessments of reliability, and that therefore they tend to over-trust or under-trust the vehicle. Conversely, the absence of a significant difference between the two values suggests that participants' trust was approximately in line with engineer reliability scores, and thus that participant trust was well calibrated.

A Bonferroni correction was applied to reduce the likelihood of Type I Errors. The corrected p-value was calculated by dividing the alpha-value ($=0.05$) by the number of scenarios (9): ($=0.05/9$) $= 0.0055$ for each scenario.

An a priori power analysis was conducted using G*Power3 (Faul et al., 2007) to test the difference from a constant value (one-sample case) using a two-tailed test, a medium effect size ($d = 0.50$), and an alpha of 0.0055. The result showed that a total sample of 59 participants was required to achieve a power of 0.80.

The one-sample Wilcoxon Signed Ranks test indicated that participants' trust was significantly lower than the engineers' reliability scores in three of the nine scenarios: Bus ($p < .001$), Crosswalk ($p < .001$), and Roadblock ($p < .001$). In these scenarios, participants underestimated the automated vehicle's capabilities. In two scenarios (Junction and Uphill), users' trust was significantly higher than the engineers' reliability scores (Junction ($p < .001$) and Uphill ($p = .003$)). Here, participants overestimated the automated vehicle. For the remaining four scenarios, the scores of the two groups did not differ significantly, suggesting that participants' trust was in line with the engineers' reliability scores (see Fig. 3 and Table 1).

3.3. Qualitative analysis

Participants' answers to open-ended questions were analysed through a software package (Atlas.ti, 2020). Their statements were marked and listed into categories (i.e., codes). These were not mutu-

ally exclusive. For example, the statement "Easy for sensors to detect the cones. Not a dangerous situation, should be fine!" (p. 4) was marked with the codes "Safe" and "Sensor". Before proceeding with the analysis, the sentences placed in each category were reviewed by a second independent observer. No major differences were found. Following this review, the frequencies of each code were normalized into percentages.

3.3.1. Engineers

When rating the reliability of the automated vehicle as not optimal (i.e., below "5"), engineers were asked to briefly explain why. Given the limited amount of data, the engineers' responses were not coded. Yet, as shown below, there was general agreement among their responses.

In several scenarios, engineer concerns were focused on object recognition: the detection of cones (Obstacle), pedestrians (Crosswalk), oncoming traffic appearing from a blind corner (Bus, Uphill), and the detection of a roadblock (Roadblock) were all mentioned as important challenges. Engineers also expressed concerns about the implementation of algorithms that would allow the automated vehicle to respect traffic rules. For this reason, scenarios Junction and Roundabout were considered particularly challenging. Conversely, Car was considered as an easy scenario to tackle.

3.3.2. Participants

The participants' explanations of their trust scores were first categorized under three super-codes: "Safe", "Unsafe", "Uncertain". "Safe", includes all statements in which participants believe that vehicle behaviour will not lead to dangerous outcomes. "Unsafe" includes statements in which participants believe that vehicle behaviour could lead to dangerous outcomes. "Uncertain" refers to statements where participants are unsure whether the vehicle will behave safely in the specific situation. The five codes "Sensor", "Steering", "Complex", "Speed" and "Visibility" were used to classify participant answers into sub-categories. These helped us understand the reasons behind their answers (see Table 2). Statements concerning object recognition are coded under the heading "Sensor". Statements concerning how the automated vehicle tackles curves are coded by "Steering". "Complex" relates to driving situations perceived as complicated for the automated vehicle. "Speed" codes for situations in which the speed of the automated vehicle is perceived as too high or too low. Finally, "Visibility" refers to the visibility of objects and upcoming traffic.

Participant answers to the open-end question show that most participants perceived the majority of the scenarios as safe (Table 2). Participants justified these assessments in terms of the clear visibility of obstacles and oncoming traffic, and the expected performance of the automated vehicle's sensors (Table 2).

None of the scenarios were thought to be clearly unsafe (Table 2). However, in all scenarios many participants expressed uncertainty about how the automated vehicle would behave (Table 2). This was expected, given that participants had never experienced an automated vehicle before and did not receive feedback on how the automated system would handle the situation from the Human-Machine-Interface (HMI) or from the experimenter. Uncertainty was strongest in situations in which obstacles and oncoming traffic were not considered to be clearly visible (i.e., Obstacle; Bus), when the detection of a crossing pedestrian (i.e., Crosswalk) or a truck (i.e., Junction) was fundamental for the safe completion of the scenario, and when the situation was considered to be very complex (i.e., Roadblock), due to the vehicle's need to perform a U-turn and find an alternative route (Table 2).

3.4. Exit questionnaire

3.4.1. Descriptive statistics

Responses to the closed items of the exit questionnaire did not show any clear preferences on the part of participants. In general,

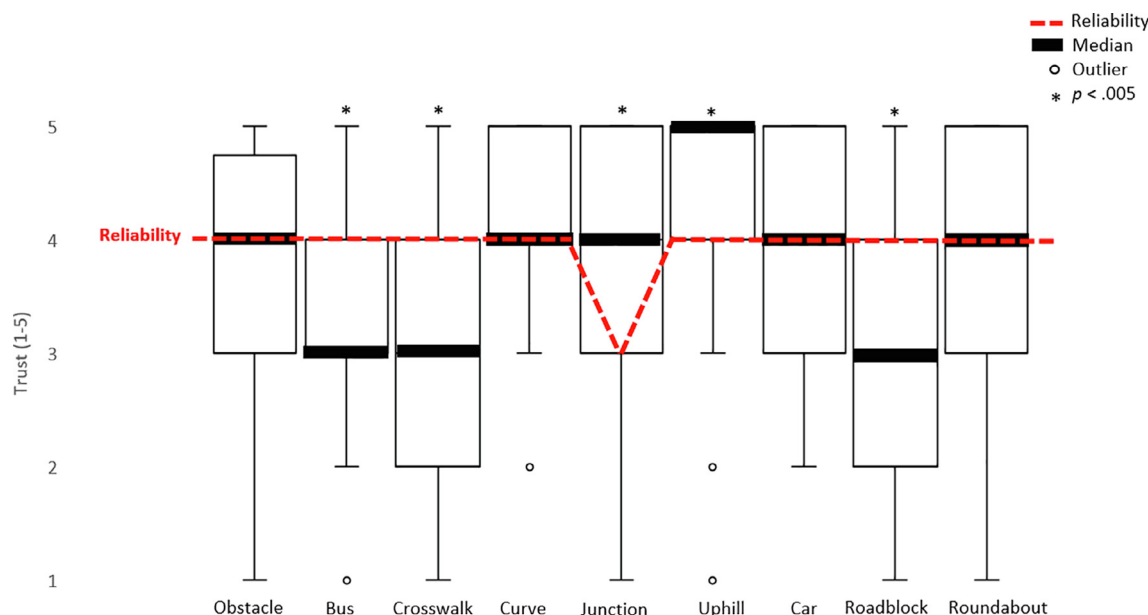


Fig. 3. Comparison between user (trust) and engineer (reliability) median scores. Lower and upper box boundaries represent 25th (quartile 1) and 75th (quartile 3) percentiles, respectively. Whiskers represent minimum and maximum reported values (excluding outliers). Data that are more than 1.5 times the interquartile range are plotted as outliers. Engineers’ reliability scores are represented through the dashed red line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Medians (\bar{x}) and standard deviations (SD) of users’ trust and engineers’ reliability scores.

Scenario	Trust	Reliability
Obstacle	$\bar{x} = 4$ SD = 1.1	$\bar{x} = 4$ SD = 0.47
Bus	$\bar{x} = 3$ SD = 1.03	$\bar{x} = 4$ SD = 0.94
Crosswalk	$\bar{x} = 3$ SD = 1.25	$\bar{x} = 4$ SD = 0.82
Curve	$\bar{x} = 4$ SD = 0.78	$\bar{x} = 4$ SD = 0.82
Junction	$\bar{x} = 4$ SD = 1.1	$\bar{x} = 3$ SD = 0.94
Uphill	$\bar{x} = 5$ SD = 0.89	$\bar{x} = 4$ SD = 0.47
Car	$\bar{x} = 4$ SD = 0.92	$\bar{x} = 4$ SD = 0.47
Roadblock	$\bar{x} = 3$ SD = 1.11	$\bar{x} = 4$ SD = 0.82
Roundabout	$\bar{x} = 4$ SD = 0.97	$\bar{x} = 4$ SD = 1.7

participants did not perceive the automated vehicle’s speed as too high or too low ($M = 3.16$, $SD = 0.45$), or the steering as being too loose or too stiff ($M = 2.77$, $SD = 1$). The size of the dashboard ($M = 3.92$, $SD = 0.95$) and the information it provided ($M = 2.27$, $SD = 1.19$) were both considered to be adequate. Nevertheless, participants would have liked to interact with a touch interface ($M = 3.69$, $SD = 1.17$). In general, automated vehicle behaviour appeared in line with participants’ expectations ($M = 3.6$, $SD = 0.97$) (see Table 3 and Appendix B).

3.4.2. Qualitative analysis

The exit questionnaire included three non-mandatory open questions. Seven codes were extrapolated from participant responses to the question “Which behaviour of the vehicle did not meet your expectations?”, and twelve from participant answers to the questions “Is there any additional information that you would have liked to be

provided with?” and “Are there any other features that you missed in the automated vehicle?” (see Table 4 and Appendix B).

The codes extrapolated from the first question revealed that, at times, the vehicle’s steering, acceleration pattern and speed were not in line with participants’ expectations. Users appeared particularly surprised by the fact that the vehicle occasionally needed to slightly readjust its position to the centre of the driving lane. Indeed, 50% of participant answers to the first open question of the exit questionnaire referred to this issue (“Steering”, Table 4). Other answers (10%) emphasized that the vehicle increased its speed (from 0 to 50 km/h) too rapidly (“Acceleration”) and that “[...] a human would increase the speed more carefully” (p. 60). Concerning the vehicle’s speed, participant statements (18%) showed that this was at times perceived as being too high (“Too fast”), although speed limits were never exceeded. Notably, in a number of occasions (11%) participants reported that the vehicle’s behaviour positively exceeded their expectations (“Better than expected”). This was underlined by statements such as “It was more humanlike than I thought, very smooth as well” (p. 49) or “I thought that I would feel uncertain and not able to trust the car while driving, but that wasn’t the case” (p. 54). To a lesser extent, participants reported the automated vehicle’s behaviour as being careless (“Careless”), appeared surprised by the fact that the vehicle kept an almost constant speed (“Constant speed”), and lacked the possibility to provide instructions to the automated system (“No communication”).

There was some overlap between the additional information that participants would have liked to have access to during the drive, and features that participants believed the automated vehicle was missing. For both questions, participant answers (35%) indicated that a visual representation of what the vehicle “sees” would be an important addition to the vehicle’s HMI (“Vehicle’s view”), and that this feature is indeed missing (21%). As stated by one of the participants: “I would like to see what the car sees and which things it detects” (p. 24). Following these lines, participants would have also liked to receive more feedback (18%) about the vehicle’s decision-making process (“Decision making”): “Some indication of the thinking process of the car. So some information on what the car will do next” (p. 55). Furthermore, participants would have liked feedback (24%) concerning the vehicle’s abil-

Table 2

Percentage of participant answers. Note: the total percentages, presented in bold, do not always correspond the sum of the single codes. This is due to the fact that even though all answers could be categorized into one of the three super-codes, participants did not always specify the elements (i.e., sub-codes) that determined their feelings of safety.

Code	Code			Scenario
	Safe	Unsafe	Uncertain	
	61%	6%	32%	
Sensor	35%	5%	10%	
Steering	0%	0%	0%	
Complex	2%	0%	2%	Obstacle
Speed	0%	0%	2%	
Visibility	24%	0%	10%	
	37%	21%	42%	
Sensor	16%	3%	15%	
Steering	0%	3%	2%	
Complex	0%	0%	5%	Bus
Speed	0%	2%	0%	
Visibility	10%	11%	15%	
	32%	35%	31%	
Sensor	16%	23%	11%	
Steering	0%	0%	0%	
Complex	0%	0%	3%	Crosswalk
Speed	0%	10%	0%	
Visibility	6%	10%	5%	
	73%	10%	18%	
Sensor	18%	2%	2%	
Steering	10%	5%	10%	
Complex	0%	0%	0%	Curve
Speed	3%	2%	3%	
Visibility	0%	2%	3%	
	61%	3%	35%	
Sensor	35%	0%	26%	
Steering	0%	0%	0%	
Complex	0%	0%	0%	Junction
Speed	2%	0%	2%	
Visibility	16%	0%	0%	
	76%	2%	23%	
Sensor	13%	0%	10%	
Steering	0%	0%	2%	
Complex	0%	0%	0%	Uphill
Speed	5%	0%	2%	
Visibility	5%	2%	5%	
	61%	6%	32%	
Sensor	13%	0%	10%	
Steering	0%	0%	2%	
Complex	0%	0%	0%	Car
Speed	5%	0%	2%	
Visibility	5%	2%	5%	
	35%	18%	47%	
Sensor	13%	5%	10%	
Steering	0%	0%	2%	
Complex	5%	6%	15%	Roadblock
Speed	0%	0%	0%	
Visibility	2%	0%	0%	
	65%	10%	26%	
Sensor	24%	5%	6%	
Steering	0%	0%	2%	
Complex	0%	3%	3%	Roundabout
Speed	0%	0%	0%	
Visibility	16%	0%	0%	

ity to handle specific situations (“Confidence indication”): “*I would like to know when the vehicle feels unable to deal with a situation and when it feels able to*”(p. 24). In line with this statement, a few answers underlined that participants lacked visual and audio alerts that would confirm the correct detection of elements that could hinder the safe completion of the driving scenario (“Confirmation”; “Alerts”). In addition, answers (9%) indicated that participants would have felt more at ease with a visual representation of the vehicle’s route (“Navigation system”) and, to a lesser extent, suggested that information could be provided through a head-up display (“Head-up display”). One participant appeared surprised that the vehicle pedals remained still when the vehicle started gaining speed (“Pedals”).

Among the other reported missing features and in line with responses to the closed items of the exit questionnaire, participant answers (17%) suggest that they lacked interaction with the vehicle (“Interaction”). This was true both in terms of “*Interaction with and information from the vehicle*” (p. 26). Concerning the information conveyed by the HMI participant answers (8%) showed that “voice” would be, for some, the preferred communication mode (“Voice”). In addition, information concerning the vehicle state (i.e., automation activated/deactivated) may be added to the vehicle’s interface (“State awareness”). One participant reported that safety could be increased by monitoring drivers’ alertness (“Monitoring alertness”).

Table 3

Descriptive statistics of the exit questionnaire. All items were rated on a 5-point Likert scale. For items 1, 2 and 3, the extremes of the scales were “Too slow (1) – Too fast (5)”, “Too lose (1)” – “Too stiff (5)” and “Sufficient (1) – Insufficient (5)”, respectively. For all the other items, the extremes were “Not at all – Extremely”.

Item	Description	Mean	Std. Deviation
1	Speed (Too slow – Too fast)	3.16	0.45
2	Steering (Too lose – Too stiff)	2.77	1
3	Dashboard info (Sufficient – Insufficient)	2.27	1.19
4	Dashboard size (Not at all – Extremely)	3.92	0.95
5	Dashboard style (Not at all – Extremely)	3.03	1.20
6	Audio (Not at all – Extremely)	3.16	1.26
7	Touch (Not at all – Extremely)	3.69	1.17
8	Voice input (Not at all – Extremely)	3.31	1.48
9	Voice communication (Not at all – Extremely)	3.19	1.36
10	Ambient light (Not at all – Extremely)	3.27	1.1
11	Human-like (Not at all – Extremely)	3.24	0.97
12	Expectations (Not at all – Extremely)	3.60	0.97

Table 4

Percentage of the codes used in the exit questionnaire.

Code	Question		
	Unexpected behaviour	Additional information	Missing features
Careless	3%		
Constant speed	3%		
Acceleration	10%		
Better than expected	10%		
No communication	5%		
Steering	50%		
Too fast	18%		
Head-up display		3%	4%
Confidence indication		24%	4%
Confirmation		9%	8%
Decision-making		18%	4%
Navigation system		9%	13%
Vehicle's view		35%	21%
Pedals		3%	4%
Alerts			8%
Monitoring alertness			4%
Interaction			17%
State awareness			4%
Voice			8%

Note: Unexpected Behaviour N = 38; Additional information N = 34; Missing Features N = 24.

4. Discussion

When testing specific driving scenarios, a mismatch emerged between the engineers' perceived reliability of the automated vehicle and the trust of our potential users. Under-trust was observed in three of the nine scenarios (i.e., Bus, Crosswalk and Roadblock). Here, crucial elements that may have hindered the safe completion of the driving situation were not immediately visible to participants. Their concerns were shared by the engineers, but to a lesser extent. Furthermore, in scenario Roadblock, participants appeared unsure of what the vehicle would do after detecting the barrier.

Over-trust was observed in two of the nine scenarios (i.e., Uphill and Junction). Engineers appeared concerned of oncoming traffic. Particularly in Junction, while most participants believed that the automated vehicle would safely cross the intersection, engineers appeared concerned about the detection of the crossing truck. In general, intersections represent a complex task for the automated system. This, as proven by the engineers' mixed responses, also applies to roundabouts.

Our results are in line with findings from several studies, showing that context-dependent characteristics of the driving scenario (e.g., road type, traffic volume, situation type) strongly impact users' trust towards automated driving systems (Frison et al., 2019; Li et al., 2019; Sonoda and Wada, 2016; Walker et al., 2018). In other words, users trust automated vehicles in some situations more than in others. This is likely due to users' perceived risks, that may change from one situation to the other and that play an important role in the development and calibration of trust in automation (Hoff and Bashir, 2015; Lee and See, 2004; Li et al., 2019; Perkins et al., 2010).

Participants' answers to the open-questions of the exit questionnaire point towards interventions that could mitigate under and over-trust and, in general, guarantee a better user experience. Their first concern was vehicle dynamics: although the automated vehicle never exceeded the 50 km/h speed limit and always kept within its lane, participants appeared concerned about its behaviour. In particular, participants did not expect that the car would need to apply small adjustments to its position. These were necessary to assure that the vehicle would be in the centre of the driving lane at all times. Furthermore, some participants perceived the vehicle speed as being too fast and its acceleration pattern as too aggressive. These results corroborate previous findings that stress the importance of the vehicle's driving style – showing how this may strongly affect driver trust and comfort (Price et al., 2016; Lee et al., 2016; Ekman et al., 2019).

An appropriate level of trust may be achieved not only by improving vehicle performance, but also by presenting to drivers information about the automated system's decisions and actions. Notably, the presentation of feedback would indirectly increase interaction with the automated driving system – something that participants reported was lacking. In this respect, participants' answers suggest that a graphical representation of the elements detected in the environment, combined with an indication of how specific driving situations may be tackled, may improve trust calibration. For example, in scenario Bus, a graphical representation of the still bus together with the path that the automated vehicle intends to follow would allow the drivers to know whether the bus has been detected by the system, and if it would be overtaken or not. Drivers' trust could then be calibrated accordingly. This suggestion is supported by Ekman et al.'s (2016) findings, showing that presenting to drivers feedback concerning the objects present on the vehicle's path increased their trust in the automated system. Furthermore, as recently pointed out by Domeyer et al. (2020), the “observability” of complex automation intentions may strongly improve human-automation interaction.

In line with these considerations, participants reported that feedback concerning the vehicle's performance would have improved their driving experience. As pointed out by Seppelt and Lee (2019), real-time feedback on the behaviour of the automated driving system promotes an accurate mental model of the system processes, and therefore may be preferred to single warnings. Indeed, authors have shown that presenting real-time feedback of the automated driving system's performance may improve drivers' trust calibration, and thus promote safer human-automation interaction (Helldin et al., 2013; Kunze et al., 2019).

The low standard deviations of participants' trust scores (see Table 1) suggest that under- and over-trust observed in our scenarios is not strongly influenced by users' individual personalities or preferences. This implies not just that corrective engineering could produce significant improvements in trust calibration, but that such changes could impact a large proportion of our potential user population. More generally, our study shows that experiencing an automated vehicle that behaves in a predictable way leads to higher trust. This result is in line with the literature, and suggests that initial trust in automation can be altered by on-road experience (Beggiato and Kreams, 2013; Endsley, 2017; Gold et al., 2015; Walker et al., 2018).

Our study has a number of limitations that should be acknowledged. The predicted reliability of the automated vehicle was assessed

by three engineers involved in the vehicle design process. Although their scores generally agreed, a larger sample size would have strengthened our results. Furthermore it may be argued that, given that road scenarios were presented to engineers and participants in a different manner, the two groups could not be truly compared. In this respect, we argue that the true knowledge of the experts could be better elicited by presenting the driving scenarios through a bird-eye view. This allowed the engineers to provide a detached judgment, uninfluenced by feelings that may come into play during a simulated drive. This is one area where reliability and trust truly differ. While the former is established through accurate knowledge concerning system performance, the latter is also influenced by feelings of uncertainty and vulnerability that come into play when experiencing the automated system. In short, although these feelings may also influence the engineers' trust, they do not affect the objective reliability of the system. Therefore, including them could have negatively affected the engineers' reliability ratings.

On a different note, although there is strong evidence for the relative ecological validity of simulator-based research (e.g., Kaptein et al., 1996; Godley et al., 2002; Meuleners and Fraser, 2015; Klüver et al., 2016), feelings of vulnerability, uncertainty and perceived risk are likely to differ in real and simulated environments. In our own study, we did not assess how "vulnerable" our participants felt during the driving experience. However, the i-CAVE Twizy is designed for deployment in situations where risk is inherently low (e.g., the university campus). Therefore, the lack of risk in our simulation likely does not affect its ecological validity for these conditions. Regarding uncertainty, users reported that they were surprised by certain aspects of the automated vehicle's behaviour (e.g., the vehicle's need to adjust its position to the centre of the lane), and that in multiple occasions they were uncertain of how the car would handle the driving task. It thus seems that the simulator induced uncertainties are similar to the ones drivers would experience on the road. Overall, many studies of trust in automated driving technology have been conducted in driving simulators (e.g., Gold et al., 2015; Hergeth et al., 2016; Molnar et al., 2018). This is mostly due to the fact that critical driving situations may lead to physical harm, and therefore cannot be safely investigated on the road. Nonetheless, more on-road research is needed to truly understand users' trust and their interactions with automated driving systems.

In line with these considerations, Li et al. (2019) recently pointed out that most studies investigating user trust towards automated vehicles are not actually measuring trust, but rather perceived vehicle reliability. We would argue that this is inevitable, since trust strongly depends on perceived vehicle reliability. As recently argued by Körber (2018), trust is an attitude, closely related to beliefs and expectations concerning the automated driving system. Therefore, trust requires an affective evaluation of the (perceived) characteristics of the automated vehicle, such as its reliability (Körber, 2018). Further evidence concerning the link between trust and perceived reliability comes from Lee and See (2004) that, in line with Mayer et al. (1995), argue that "performance" is a crucial dimension of trust in automation. "Performance" includes the system's perceived reliability, competency and ability to perform certain tasks.

In conclusion, although the engineers consistently gave positive assessments of the reliability of the vehicle, it should be stressed that since the automated vehicle is not yet road-test ready, its objective reliability is currently unknown. The goal of this study was not to test whether participant views concerning vehicle behaviour were objectively correct, but to explore their alignment with engineer evaluations

– the best information available before actual road testing. In this respect, our study shows that user trust and engineer evaluations of vehicle reliability are often misaligned, and points towards solutions that may lead to calibrated trust.

Participants' suggestions will be discussed with our engineering team, implemented in an updated simulated version of the Twizy, and tested with a new pool of users. Overall, the adoption of our methods, or similar methods, can make a significant contribution to safety and usability of future automated vehicles. In this respect, a user-driven approach – such as the one described in this manuscript – allows the implementation and investigation of user suggestions in an early stage of development.

5. Conclusion

Our study shows that users' trust and engineers' evaluations of vehicle reliability are often misaligned. Such misalignment may be mitigated by acting on the vehicle's dynamics and on its HMI. Concerning vehicle behaviour, our results suggest that first-time users may prefer an automated vehicle with a more defensive driving style. Therefore, a vehicle that keeps a more centred position in its driving lane, drives more slowly and avoids heavy accelerations. As concerns the HMI, our findings suggest that a visual representation of the objects detected by the automated driving system in its surroundings, combined with real-time feedback on vehicle performance, could improve trust calibration. Overall, our results show that – before actual road testing – the comparison of engineer perceptions of reliability and user trust can lead to important suggestions for the improvement of vehicle design.

6. Data availability statement

The data collected during the study are available at <http://doi.org/10.17605/OSF.IO/DE8KM>.

CRedit authorship contribution statement

F. Walker: Conceptualization, Methodology, Software, Validation, Data curation, Formal analysis, Writing - original draft, Writing - review & editing, Visualization, Project administration. **J. Steinke:** Methodology, Software, Formal analysis, Investigation, Data curation, Writing - review & editing. **M.H. Martens:** Resources, Writing - review & editing, Supervision, Funding acquisition. **W.B. Verwey:** Resources, Writing - review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

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

Acknowledgments

This research is supported by the Dutch Domain Applied and Engineering Sciences, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs (project number 14896). The authors would like to thank the engineers of the i-CAVE project that took part in this study.

Appendix A: Driving scenarios

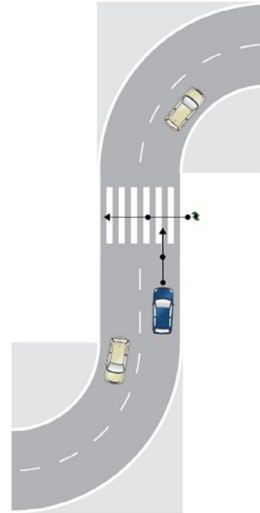
Boxes

Three traffic cones are blocking the way. These should be avoided by passing on the left. There is no oncoming traffic.



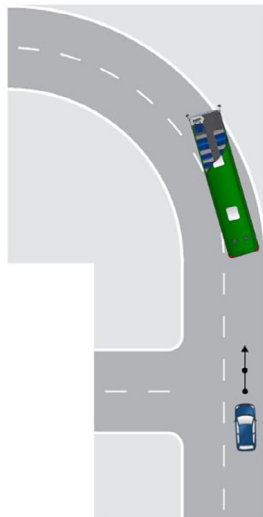
Crosswalk

A pedestrian is standing at a crosswalk and wants to cross the road. The vehicle should stop and wait until the pedestrian has crossed the road.



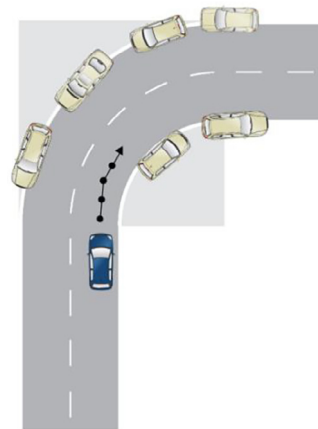
Bus

The bus ahead is standing still and should be overtaken. There is no oncoming traffic



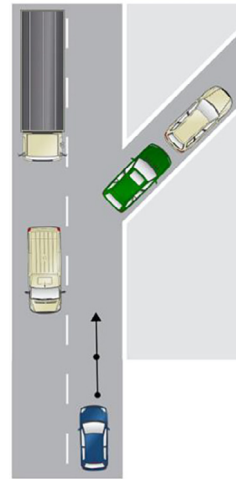
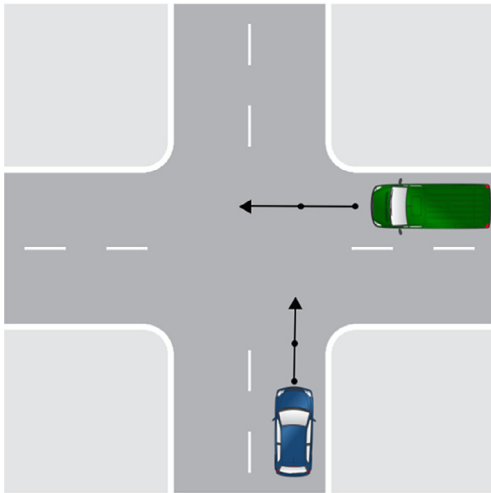
Curve

The vehicle takes a curve. There are vehicles parked on the left and right hand side of the curve.



Junction

A truck, approaching from the right, has the right of way. The vehicle needs to stop to let the truck pass.

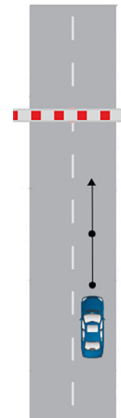
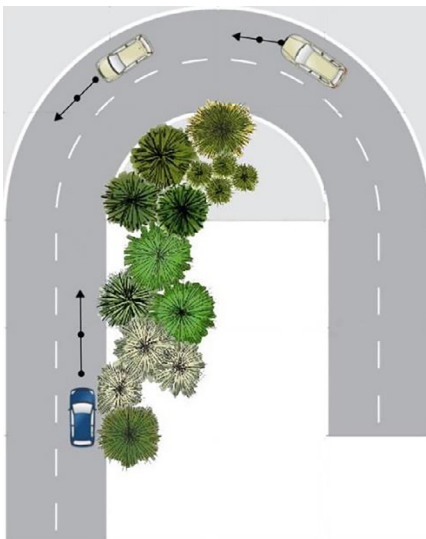


Roadblock

The road is closed entirely. The vehicle cannot continue on this road.

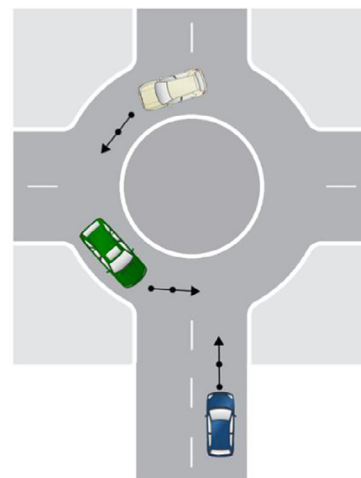
Uphill

The automated vehicle is approaching a blind curve (due to bushes on the right hand side of the road). The road is uphill. Oncoming traffic comes around the curve.



Roundabout

The automated vehicle should enter the roundabout. The roundabout is busy with oncoming traffic.



Car

The car ahead has left your lane, but not entirely. The rear of the car is still on your lane. There is oncoming traffic.

Appendix B: Exit questionnaire

- The speed of the vehicle was
Too slow Too fast
- The steering of the vehicle was
Too loose Too stiff
- The information provided by the dashboard was
Sufficient Insufficient
- Was the size of the dashboard sufficient?
Not at all Extremely
- Do you prefer a digitally styled dashboard over an analogue style?
Not at all Extremely
- Would you like audio information?
Not at all Extremely
- Would you like to be able to interact with the car via a touch interface?
Not at all Extremely
- Would you like to be able to interact with the car via voice input?
Not at all Extremely
- Would you like the vehicle to communicate to you with a voice?
Not at all Extremely
- Would you like ambient lighting to provide information inside the car?
Not at all Extremely
- Was the behaviour of the car humanlike?
Not at all Extremely
- Was the behaviour of the vehicle in line with your expectations?
Not at all Extremely
- If not, which behaviour of the vehicle did not meet your expectations?
Is there any additional information that you would have liked to be provided with?
Are there any other features that you missed in the automated vehicle?

References

- Atlas.ti [website], 2020. Retrieved from <https://atlasti.com/product/what-is-atlas-ti/> (accessed 15 April, 2020).
- Beggiato, M., Krems, J.F., 2013. The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transp. Res. Part F: Traffic Psychol. Behav.* 18, 47–57. <https://doi.org/10.1016/j.trf.2012.12.006>.
- Brower, H.H., Schoorman, F.D., Tan, H.H., 2000. A model of relational leadership. *Leadership Q.* 11 (2), 227–250. [https://doi.org/10.1016/S1048-9843\(00\)00040-0](https://doi.org/10.1016/S1048-9843(00)00040-0).
- Carsten, O., Martens, M.H., 2019. How can humans understand their automated cars? HMI principles, problems and solutions. *Cogn. Tech. Work* 21 (1), 3–20. <https://doi.org/10.1007/s10111-018-0484-0>.
- Choi, J.K., Ji, Y.G., 2015. Investigating the importance of trust on adopting an autonomous vehicle. *Int. J. Human-Computer Interaction* 31 (10), 692–702. <https://doi.org/10.1080/10447318.2015.1070549>.
- Domeyer, J.E., Lee, J.D., Toyoda, H., 2020. Vehicle automation—other road user communication and coordination: theory and mechanisms. *IEEE Access* 8, 19860–19872. <https://doi.org/10.1109/ACCESS.2020.2969233>.
- Ekman, F., Johansson, M., Bligård, L.O., Karlsson, M., Strömberg, H., 2019. Exploring automated vehicle driving styles as a source of trust information. *Transp. Res. Part F: Traffic Psychol. Behav.* 65, 268–279. <https://doi.org/10.1016/j.trf.2019.07.026>.
- Ekman, F., Johansson, M., Sochor, J., 2016. To See or Not to See: The Effect of Object Recognition on Users' Trust in "Automated Vehicles". In: Proceedings of the 9th Nordic Conference on Human-Computer Interaction, pp. 1–4. <https://doi.org/10.1145/2971485.2971551>.
- Endsley, M.R., 2017. Autonomous driving systems: a preliminary naturalistic study of the tesla model S. *J. Cognitiv Eng. Decision Making* 11 (3), 225–238. <https://doi.org/10.1177/1555343417695197>.
- Fagnant, D.J., Kockelman, K., 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transp. Res. Part A: Policy Practice* 77, 167–181. <https://doi.org/10.1016/j.tra.2015.04.003>.
- Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39 (2), 175–191. <https://doi.org/10.3758/BF03193146>.
- Frison, A.K., Wintersberger, P., Liu, T., Riener, A., 2019. Why do you like to drive automated? a context-dependent analysis of highly automated driving to elaborate requirements for intelligent user interfaces. In: Proceedings of the 24th International Conference on Intelligent User Interfaces, pp. 528–537. <https://doi.org/10.1145/3301275.3302331>.
- Ghazizadeh, M., Lee, J.D., Boyle, L.N., 2012. Extending the technology acceptance model to assess automation. *Cogn. Tech. Work* 14 (1), 39–49. <https://doi.org/10.1007/s10111-011-0194-3>.
- Godley, S.T., Triggs, T.J., Fildes, B.N., 2002. Driving simulator validation for speed research. *Accid. Anal. Prev.* 34 (5), 589–600. [https://doi.org/10.1016/S0001-4575\(01\)00056-2](https://doi.org/10.1016/S0001-4575(01)00056-2).
- 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 Manuf.* 3, 3025–3032. <https://doi.org/10.1016/j.promfg.2015.07.847>.
- Hellidin, T., Falkman, G., Riveiro, M., Davidsson, S., 2013. Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. In: Proceedings of the 5th international conference on automotive user interfaces and interactive vehicular applications, pp. 210–217. <https://doi.org/10.1145/2516540.2516554>.
- Hergeth, S., Lorenz, L., Vilimek, R., Krems, J.F., 2016. Keep your scanners peeled: gaze behavior as a measure of automation trust during highly automated driving. *Hum. Factors* 58 (3), 509–519. <https://doi.org/10.1177/0018720815625744>.
- Hoff, K.A., Bashir, M., 2015. Trust in automation: integrating empirical evidence on factors that influence trust. *Hum. Factors* 57 (3), 407–434. <https://doi.org/10.1177/0018720814547570>.
- i-CAVE [website], 2020. Retrieved from <https://i-cave.nl/> (accessed 15 April 2020).
- Jian, J.-Y., Bisantz, A.M., Drury, C.G., 2000. Foundations for an empirically determined scale of trust in automated systems. *Int. J. Cognitive Ergonomics* 4 (1), 53–71. https://doi.org/10.1207/S15327566JCE0401_04.
- Kaptein, N.A., Theeuwes, J., van der Horst, R., 1996. Driving simulator validity: some considerations. *Transp. Res. Rec.* 1550 (1), 30–36. <https://doi.org/10.1177/0361198196155000105>.
- Khastgir, S., Birrell, S., Dhadyalla, G., Jennings, P., 2017. Calibrating trust to increase the use of automated systems in a vehicle. *Adv. Intell. Syst. Comput.* 484, 535–546. https://doi.org/10.1007/978-3-319-41682-3_45.
- Klüver, M., Herrigel, C., Heinrich, C., Schöner, H.-P., Hecht, H., 2016. The behavioral validity of dual-task driving performance in fixed and moving base driving simulators. *Transp. Res. Part F: Traffic Psychol. Behav.* 37, 78–96. <https://doi.org/10.1016/j.trf.2015.12.005>.
- Körber, M., 2018. Theoretical considerations and development of a questionnaire to measure trust in automation. In: Proceedings of the 20th Congress of the International Ergonomics Association, pp. 13–30. https://doi.org/10.1007/978-3-319-96074-6_2.
- Kunze, A., Summerskill, S.J., Marshall, R., Filtness, A.J., 2019. Automation transparency: implications of uncertainty communication for human-automation interaction and interfaces. *Ergonomics* 62 (3), 345–360. <https://doi.org/10.1080/00140139.2018.1547842>.
- Kyriakidis, M., de Winter, J.C.F., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., Martens, M.H., Bengler, K., Andersson, J., Merat, N., Reed, N., Flament, M., Hagenzieker, M., Happee, R., 2019. A human factors perspective on automated driving. *Theor. Issues Ergonomics Sci.* 20 (3), 223–249. <https://doi.org/10.1080/1463922X.2017.1293187>.
- Lee, J., Kim, N., Imm, C., Kim, B., Yi, K., Kim, J., 2016. A question of trust: An ethnographic study of automated cars on real roads. In: Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 201–208. <https://doi.org/10.1145/3003715.3005405>.
- Lee, J.D., See, K.A., 2004. Trust in automation: designing for appropriate reliance. *Human Factors: J. Human Factors Ergonomics Society* 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: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 177–185. <https://doi.org/10.1145/3342197.3344525>.
- Litman, T., 2017. Autonomous vehicle implementation predictions (p. 28). Victoria, Canada: Victoria Transport Policy Institute. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.640.2382&rep=rep1&type=pdf>.
- Mayer, R.C., Davis, J.H., Schoorman, F.D., 1995. An integrative model of organizational trust. *AMR* 20 (3), 709–734. <https://doi.org/10.5465/amr.1995.9508080335>.
- Meuleners, L., Fraser, M., 2015. A validation study of driving errors using a driving simulator. *Transp. Res. Part F: Traffic Psychol. Behav.* 29, 14–21. <https://doi.org/10.1016/j.trf.2014.11.009>.
- Milakis, D., van Arem, B., van Wee, B., 2017. Policy and society related implications of automated driving: a review of literature and directions for future research. *J. Intell. Transp. Syst.* 21 (4), 324–348. <https://doi.org/10.1080/15472450.2017.1291351>.
- Molnar, L.J., Ryan, L.H., Pradhan, A.K., Eby, D.W., St. Louis, R.M., Zakrajsek, J.S., 2018. Understanding trust and acceptance of automated vehicles: An exploratory simulator study of transfer of control between automated and manual driving. *Transp. Res. Part F: Traffic Psychol. Behav.* 58, 319–328. <https://doi.org/10.1016/j.trf.2018.06.004>.
- Muir, B.M., 1987. Trust between humans and machines, and the design of decision aids. *Int. J. Man Mach. Stud.* 27 (5-6), 527–539. [https://doi.org/10.1016/S0020-7373\(87\)80013-5](https://doi.org/10.1016/S0020-7373(87)80013-5).
- Nyhan, R.C., 2000. Changing the paradigm: trust and its role in public sector organizations. *Am. Rev. Public Administration* 30 (1), 87–109. <https://doi.org/10.1177/02750740022064560>.
- Parasuraman, R., Riley, V., 1997. Humans and automation: use, misuse, disuse, abuse. *Hum. Factors* 39 (2), 230–253. <https://doi.org/10.1518/00187209778543886>.
- Parasuraman, R., Sheridan, T.B., Wickens, C.D., 2008. Situation awareness, mental workload, and trust in automation: viable, empirically supported cognitive

- engineering constructs. *J. Cognitive Eng. Decision Making* 2 (2), 140–160. <https://doi.org/10.1518/155534308X284417>.
- Payre, W., Cestac, J., Delhomme, P., 2016. Fully Automated Driving: Impact of Trust and Practice on Manual Control Recovery. *Hum. Factors* 58 (2), 229–241. <https://doi.org/10.1177/0018720815612319>.
- Perkins, LeeAnn, Miller, J.E., Hashemi, A., Burns, G., 2010. Designing for human-centered systems: situational risk as a factor of trust in automation. *Proc. Human Factors Ergonomics Society Annual Meeting* 54 (25), 2130–2134. <https://doi.org/10.1177/154193121005402502>.
- Price, M.A., Venkatraman, V., Gibson, M., Lee, J., Mutlu, B., 2016. Psychophysics of trust in vehicle control algorithms (No. 2016-01-0144). *SAE Technical Paper*. <https://doi.org/10.4271/2016-01-0144>.
- SAE, 2018. (R) Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems, 1-35. https://doi.org/10.4271/J3016_201806.
- Seppelt, B.D., Lee, J.D., 2019. Keeping the driver in the loop: Dynamic feedback to support appropriate use of imperfect vehicle control automation. *Int. J. Hum Comput Stud.* 125, 66–80. <https://doi.org/10.1016/j.ijhcs.2018.12.009>.
- Shapiro, S.P., 1987. The Social control of impersonal trust. *Am. J. Sociol.* 93 (3), 623–658. <https://doi.org/10.1086/228791>.
- Sonoda, K., Wada, T., 2016. Driver's trust in automated driving when sharing of spatial awareness. In: *Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 002516–002520. <https://doi.org/10.1109/SMC.2016.7844618>.
- Urmson, C., Whittaker, W., 2008. Self-driving cars and the urban challenge. *IEEE Intell. Syst.* 23 (2), 66–68. <https://doi.org/10.1109/MIS.2008.34>.
- Verberne, F.M.F., Ham, J., Midden, C.J.H., 2012. Trust in smart systems: sharing driving goals and giving information to increase trustworthiness and acceptability of smart systems in cars. *Hum. Factors* 54 (5), 799–810. <https://doi.org/10.1177/0018720812443825>.
- Walker, F., Boelhouwer, A., Alkim, T., Verwey, W.B., Martens, M.H., 2018. Changes in trust after driving level 2 automated cars. *J. Adv. Transp.* 2018, 1–9. <https://doi.org/10.1155/2018/1045186>.
- Walker, F., Wang, J., Martens, M.H., Verwey, W.B., 2019. Gaze behaviour and electrodermal activity: Objective measures of drivers' trust in automated vehicles. *Transp. Res. Part F: Traffic Psychol. Behav.* 64, 401–412. <https://doi.org/10.1016/j.trf.2019.05.021>.
- Waytz, A., Heafner, J., Epley, N., 2014. The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *J. Exp. Soc. Psychol.* 52, 113–117. <https://doi.org/10.1016/j.jesp.2014.01.005>.
- Wivw GmbH-Silab, 2018. *Driving Simulation and SILAB [company website] [online]*. Retrieved from <https://wivw.de/en/silab> (accessed 15 April 2020).