Hindawi
Journal of Advanced Transportation
Volume 2018, Article ID 1045186, 9 pages
https://doi.org/10.1155/2018/1045186

Research Article

Changes in Trust after Driving Level 2 Automated Cars

Francesco Walker [1], Anika Boelhouwer,[1] Tom Alkim,[2] Willem B. Verwey,[1] and Marieke H. Martens[1]

[1] *University of Twente, Netherlands*
[2] *Rijkswaterstaat, Ministry of Infrastructure and Water Management, Netherlands*

Correspondence should be addressed to Francesco Walker; f.walker@utwente.nl

Received 21 December 2017; Accepted 15 May 2018; Published 5 August 2018

Academic Editor: Jay Katupitiya

Overtrust and undertrust are major issues with partially automated vehicles. Ideally, trust should be *calibrated* ensuring that drivers' subjective feelings of safety match the objective reliability of the vehicle. In the present study, we examined if drivers' trust toward Level 2 cars changed after on-road experience. Drivers' self-reported trust was assessed three times: before having experience with these vehicles, immediately after driving two types of vehicles, and two weeks after the driving experience. Analysis of the results showed major changes in trust scores after the on-road driving experience. Before experiencing the vehicles, participants tended to overestimate the vehicle capabilities. Afterwards they had a better understanding of vehicles' limitations, resulting in better calibrated trust.

1. Introduction

Advanced driver assistance systems (ADAS) are technologies that support the human driver in the driving task. This is done by providing information and warnings, and automating demanding and repetitive tasks [1–3]. Several modern vehicles are equipped with a combination of ADAS. Examples include Volvo's Adaptive Cruise Control with Steer Assistance, BMW's Traffic Jam Assistant, Toyota's Automated Driving Highway Assistant, and Tesla's Autopilot [4]. The Society of Automotive Engineers has formally classified these vehicles as "Level 2", or vehicles equipped with "partial automation" [5]. With SAE Level 2, or partial automation, the driver is still responsible for deciding when and whether to engage the automated system. Furthermore, with these systems, the driver should always monitor the driving environment and be ready to respond to system failures [5].

The benefits of partial automation will not be realized until Level 2 vehicles are widely accepted and used on the road [6]. If drivers do not trust the automated system, it is likely that they will overrule the vehicle's decisions or not activate the partially automated system at all. This will reduce the potential positive safety impact of the technology [7]. In their extension of the Technology Acceptance Model (TAM), Choi and Ji [8] showed that perceived usefulness and intention to use the technology were strongly mediated by trust [8]. A similar conclusion was reached by Ghazizadeh and colleagues [9]. Through their Automation Acceptance Model, the authors suggested that trust is, in fact, one of the main factors affecting users' acceptance of automation technology.

In line with these considerations, reports show that although driver awareness of automation technology is increasing, usage is still limited [10]: A survey conducted in 2017 by the news agency AAA showed that while drivers of the United States look for autonomous technologies in their next car, only 25% would feel safe actually driving an autonomous vehicle. Despite these results, it is also possible for drivers to overestimate the capabilities of the vehicle, relying excessively on SAE Level 2 functions even under unsafe conditions or outside of the Operational Design Domain (ODD). With partially self-driving (i.e., Level 2) cars, as with other automated systems, the goal of human factors research should NOT be to maximize drivers' trust, but to *align drivers' subjective perceptions of safety with the actual reliability of the vehicle*, thus avoiding over- and underestimation of the car's capabilities (e.g., [11, 12]). This is commonly known as "trust calibration" (e.g., [11, 13]). Notably, trust calibration cannot be achieved if drivers are unable to discriminate situations in which the vehicle will behave reliably from situations in which it will not [14]. In other words, to align drivers' subjective feelings of safety with the actual reliability of the vehicle, drivers need to know how the system behaves in a variety of situations. The ability to discriminate between different situations is commonly known as "resolution" [14].

In general, paraphrasing Ghazizadeh and colleagues, trust affects use, and use affects trust [9]. Importantly, the relationship between trust and use is not always positive: Drivers might not trust the technology but still use it, and vice versa. Furthermore, the relationship between trust and use is mediated by drivers' initial trust levels and the objective reliability of the technology. Thus, calibrating drivers' trust in Level 2 vehicles requires a better understanding of their initial trust toward the vehicles, and how this is influenced by vehicle experience. Most of what we know about trust in automation comes from studies of automated aircraft or production systems [6]. These studies indicate that user trust in automation is strongly affected by system performance, unexpected failures of the automated system [15], and predictability of system behaviour [6, 11].

In the automotive domain, several studies have analysed drivers' trust in relation to Adaptive Cruise Control (ACC). ACC is an Advanced Driving Assistance System that automatically regulates speed and headway distance. Rajaonah and colleagues [16] used a simulated vehicle equipped with ACC to analyse driver behaviour when a truck suddenly cut into the driver's lane. Trust was measured via self-reports. The study identified two groups of drivers: One group that started braking before the truck pulled into their lane—correct anticipatory behaviour—and a second group that started braking only when the truck was already in their lane. Given that drivers should not rely on ACC during hazardous situations, the latter behaviour suggests overtrust in the system [16]. The group that correctly disengaged the ACC before the cut-in event reported high trust in their interaction with the ACC system, but not in the system itself. This result suggests that this group had a clear understanding of the ACC's limitations, which may allow appropriate trust calibration [16]. A survey conducted by Dickie and Boyle [17] showed that drivers who are unaware of the way ACC works were unsure about how to use it, tending to overestimate the capabilities of the system. The authors argue that the inappropriate expectations of "unaware" and "unsure" drivers lead to poor trust calibration (see also [13]).

A driving simulator study by Gold and colleagues [6] suggests that driver experience with SAE Level 3 vehicles can influence drivers' trust in the system. In such vehicles the system, and not the driver, is responsible for monitoring the driving environment [5]. Thus, in Level 3 vehicles drivers can perform all sorts of activities while the car takes care of the driving task. In Gold et al.'s [6] study participants had been briefed that the vehicle could cope with most driving scenarios. Thus, monitoring the system was unnecessary. Nevertheless, participants were also told that the vehicle could not cope with every driving scenario and that the system would emit a warning signal when it could not cope. Participants had seven seconds to take back control of the vehicle. Drivers experienced three take-over scenarios during a twenty-minute simulated drive. Interestingly, drivers' clearer understanding of the vehicles' limitations (i.e., take-over scenarios) due to on-hand experience leads to a decreased perception of safety advantages, but also to an increase in trust. In line with these results, a study investigating drivers' experience with Forward Collision Warnings showed that a learning period helped drivers to achieve a better understanding of system behaviour, ultimately improving their interactions with the system and increasing their trust [18].

Overall, the literature suggests that trust strongly affects how drivers interact with SAE Level 2 technology and that a clear understanding of vehicles' limitations is important for trust calibration [6, 13, 16–19]. As pointed out by Muir [20], trust can be built by giving to the user the chance to experience the system in a variety of situations. Notably, by doing so, on-road experience with real Level 2 vehicles can increase drivers' resolution [14]. Improving this ability could play a key role in calibrating drivers' trust.

Previous work on drivers' trust in automation has been mostly carried out in driving simulators and analysed the effect of specific ADAS (e.g., ACC, Forward Collision Warning System) or, as in the case of Gold et al. [6], systems that were not yet available on the market. They thus provide no information about drivers' trust after experiencing the systems on real roads, during a variety of conditions, and when driving different vehicles equipped with similar automated functions during the same driving session.

In our study, we explored how on-road experience with Level 2 vehicles influences drivers' trust in automation. In line with Gold et al. [6] and Koustanaï et al.'s [18] results, we expected that real-life experience would give drivers a better understanding of system behaviour, thereby improving trust calibration. In the study, therefore, participants were asked to fill in a trust questionnaire before, immediately after, and two weeks after driving two different Level 2 vehicles on motorway, urban and rural roads. The impact of the driving experience on participants' trust was assessed through the comparison of the pre- and post-trust scores. The "2-week" measurement was included to measure if changes in trust would fade after two weeks. Features such as age, gender, and views of new technology were also taken into account. To our knowledge, this is the first time trust scores have been collected before and after on-road experience with real Level 2 vehicles.

2. Method

2.1. Participants. Participants in the study were all employees of Rijkswaterstaat, the Dutch Ministry of Infrastructure and Water Management. None of the participants had driven or been driven (as a passenger) in a Level 2 vehicle before joining the study. Drivers were asked to fill in the same questionnaire three times (see "Procedure" section). Only participants who completed all three questionnaires (i.e., measurement 0, measurement 1, and measurement 2) were included in the analysis. This resulted in a sample of 106

Figure 1: Picture following description of scenario 1.

Figure 2: Picture following description of scenario 2.

participants of which 77.4% (N = 82) were male. This group had a mean age of 43.77 years (SD = 10.14). The youngest participant was 24 years old and the oldest 62.

2.2. Questionnaire. The questionnaire was composed of 34 items. Of these, 4 concerned participants' age, gender, travelling profile, and attitudes toward new technology. A subsection of the questionnaire, composed of 12 items, investigated drivers' trust toward Level 2 cars in twelve different scenarios. This was the focus of the current study.

In this section of the questionnaire, participants were first asked to indicate which driving scenarios they thought partially automated cars could or could not handle. Then, after experiencing the vehicles, participants were asked to imagine driving a Level 2 car in these different scenarios. More specifically, drivers were asked to indicate to what extent they *trusted* the system (i.e., the Level 2 car) to cope with each scenario. This was done through the statement "I trust that the system can handle the situation, thus, I do not need to take over control". Drivers were asked to indicate, through a 5-point Likert scale (i.e., 1 indicating "I fully disagree" and 5 indicating "I fully agree"), how much they agreed with the statement. Each of the 12 items was accompanied by an image illustrating the scenario (e.g., Figures 1 and 2). Scenarios were presented in Dutch and translated into English for this manuscript. The 12 scenarios were as follows.

(1) You are driving on a highway with an average amount of traffic. The system is set to automatically keep distance from the vehicle in front, and to keep within the lane (see Figure 1).

(2) You drive on the highway and the vehicle in front brakes hard (see Figure 2).

(3) A car overtakes you in the right lane.

(4) You drive on the highway with a lot of rain, so the visibility is poor.

(5) You drive in a work area with yellow road marking in addition to the normal white road marking.

(6) A car merges in from the right lane and has a much lower speed than yours.

(7) A deer suddenly crosses the highway – a situation in which normally you would brake.

(8) The lane where you drive ends, you need to move to the lane to your right.

(9) You are approaching a curve on the highway.

(10) You are driving on the highway behind a motorcycle instead of behind a car.

(11) You are driving on the highway and the speed limit decreases. Your speed should adapt to the new speed limit.

(12) You are driving at night on the highway.

2.3. Vehicles. The following Level 2 vehicles, all 2016 models, were selected for the study: Two Teslas (a Model S and a Model X), two Mercedes (an E-200 and an E-350-E), and two Volvos (a XC90 and a V90). All the vehicles were equipped with similar SAE Level 2 functions. More detailed information on the vehicle features described in this section can be retrieved from Tesla, Mercedes, and Volvo official websites.

2.3.1. Tesla Model S and Model X. Tesla's combination of Advanced Driving Assistance Systems (ADAS) falls under the name of *Autopilot*. Tesla's Autopilot features Automatic Emergency Braking (AEB), Lane Keeping capabilities (i.e., "Autosteer"), and ACC. Autopilot also allows the vehicle to automatically change lanes when prompted by the driver's tap on the turn signal lever. The 2016 version of Tesla's Autopilot could already detect pedestrians, a feature later significantly improved through the v8.0 software update.

2.3.2. Mercedes E-200 and E-350-E. Mercedes's combination of ADAS is known as *Drive Pilot*. Drive Pilot features Active Break Assist, ACC (i.e., "Distronic Cruise Control"), and Lane Keeping functions (i.e., "Active Lane Keeping Assist"). Other ADAS include Speed Limit Pilot. This is a subfunction of Drive Pilot that detects speed limits and adjusts the vehicle's speed accordingly. Like the Tesla, the Mercedes E class comes with Active Lane Change Assist, allowing the driver to initiate an automatic lane change by lightly touching the turn control. Drive Pilot does not provide pedestrian detection. Pedestrian and animal recognition was implemented in the 2017 models through the *Intelligent Drive* software update. However, the study reported here only used 2016 models.

2.3.3. Volvo XC90 and V90. *Pilot Assist* is the name chosen by Volvo to describe the combination of ADAS in their XC90 and V90 2016 models. Pilot Assist features ACC, AEB, and Lane Keeping functions. Volvo vehicle user manuals clearly specify that Pilot Assist is not a collision warning system: Even if Pilot Assist can clearly detect leading vehicles, it cannot detect pedestrians, cyclists, animals, and motorcycles. *Pilot Assist II*, implemented in the Volvo XC90 2017 model (thus, not available for this study), features pedestrian, cyclist, and Large Animal Detection functions. Unlike Tesla's Autopilot and Mercedes's Drive Pilot, Volvo's Pilot Assist does not support lane changing.

2.4. Procedure. Five separate driving sessions were organised. Each participant joined only one of the five sessions. Participants were asked to fill in the same questionnaire (see "Questionnaire" section) three times: before the session ("measurement 0"); immediately after the session ("measurement 1"); and two weeks after the session ("measurement 2"). In each session, participants drove and were also passengers in two of the Level 2 vehicles described earlier (see "Vehicles" section). Each vehicle carried a driver, two passengers in the back seats (also participants in the study), and a PRO DRIVE expert or car dealer expert in the front passenger seat. The expert explained the functionalities of the vehicle and made sure that overreliance in automation would not lead to dangerous situations. The expert also prompted participants to use the vehicle's Level 2 functions. In each driving session, participants drove for ~ 20 minutes on a predefined route including a motorway, an urban road, and a rural road.

2.5. Analysis. A Bonferroni correction was applied to reduce the likelihood of Type I Errors. The corrected p value was calculated by dividing the alpha-value (α_{original} = .05) by the number of scenarios (12): (α_{altered} = .05/12) =.004 for each scenario. A Friedman test was used to assess differences between measurement 0 (m0), measurement 1 (m1), and measurement 2 (m2). If differences between the three measurements were found, a Wilcoxon Signed Ranks test was used to assess differences between m0 and m1, m0 and m2, and m1 and m2. If the Friedman test showed differences between the three measurements, a Mann–Whitney U test was performed to analyse the effect of gender. A Kruskal-Wallis test was used to analyse innovation and age effects.

2.6. Exclusions. Participants who did not fill in the questionnaire at all three time points (i.e., m0, m1, and m2) were excluded from the analysis. 60 participants were excluded on these grounds. Missing values for any of the 12 items were replaced by the mean of the other participants' responses. Missing values were only present at m1. In no case were there more than four missing values for a single scenario.

3. Results

3.1. Participants. None of the participants selected for this study had ever experienced a Level 2 car, as a driver or a passenger. Of the final sample of 106 participants, 46.7% of the drivers reported that they used their car more often than public transport, 37.1% reported that they used public transport more than the car, and 16.2% reported equal use. None of the respondents reported exclusive use of public transport or the car. Questioned about their attitudes to new technology, 1.90% viewed themselves as innovators, 20.8% as early adopters, 48.10% as early majority, 28.30% as late majority, and 0.9% as laggards. For the purposes of the subsequent analysis, these categories were condensed into three innovation profiles: "innovators/early adopters", "early majority", and "late majority/laggards". A Mann–Whitney test showed a significant main effect of gender on innovation profile, with men considering themselves more innovative than women (U = 736.50; p = .043). Ages were categorized into three groups: <35, 35-50, 50+. Questionnaire responses showed no significant effect of age, gender, or innovation profile (data not shown).

3.2. Responses to Questionnaire. For seven of the twelve scenarios, Wilcoxon Signed Ranks tests showed significant differences (i.e., $p <= .004$) between m0 and m1 (see Figure 3 and Table 1). In scenarios 1 and 2 (see Table 1), trust *increased* after vehicle experience. All the Level 2 vehicles included in this study could cope with scenarios 1 and 2. In scenario 1 (ACC and Lane Keeping), the high scores reported at m0 (M = 4.21) showed relatively high trust for the ACC and Lane Keeping functions of the systems. The increase in trust reported at m1 (M = 4.62) suggests that the vehicles' ACC and Lane Keeping functions exceeded participants' expectations. Drivers' responses to scenario 2 (brake preceding vehicle) suggest that experiencing this scenario also had a positive impact on participants' trust ratings.

In scenarios 4, 7, 8, 9, and 12, Wilcoxon Signed Ranks tests showed significant *decreases* in drivers' trust after vehicle experience (see Figure 3 and Table 1). In scenario 4 (rain), drivers' trust decreased after they had driven the vehicles, suggesting an initial overestimation of the vehicles' capabilities. Bad weather still represents a challenge for Level 2 technology, for instance, because it makes it more difficult to detect road markings. Even though participants did not directly experience this scenario, the drop in trust from m0 (M = 2.96) to m1 (M = 2.57) can probably be explained by a clearer understanding of vehicles' limitations. Scenario 7 (crossing deer) implies that the vehicle should automatically brake in response to a crossing deer. Animal detection currently represents a big challenge for automakers, and none of the 2016 vehicles selected for the study were equipped with this feature. Even though none of the drivers experienced the scenario on the road, the reported decrease in trust from m0 (M = 2.69) to m1 (M = 2.20) suggests a better understanding of the systems' limitations after the driving experience. None of the Level 2 vehicles were capable of handling scenario 8 (merging from left lane). As with scenario 7, drivers' decrease in trust from the pre- to the postmeasurements suggests an initial overestimation of the vehicles' capabilities, but also a clearer understanding of the systems' limitations after on-road experience with the vehicles. In scenario 9 (curve on motorway), even when Lane Keeping functions were activated, the selected Level 2 vehicles would occasionally stray out of lane when facing tight curves. Most drivers experienced this limitation, and their trust decreased from m0 (M = 3.63) to m1 (M = 2.81). Drivers did not have the chance to experience scenario 12 (driving at night). Here, drivers' trust decreased from the pre- (M = 3.61) to the post- (M = 3.28) measurement. In scenario 10 (driving behind a motorbike), no significant difference was found between m0 and m1. However, the Wilcoxon Signed Ranks test showed a significant difference between m0 (M = 3.5) and m2 (M = 3.15).

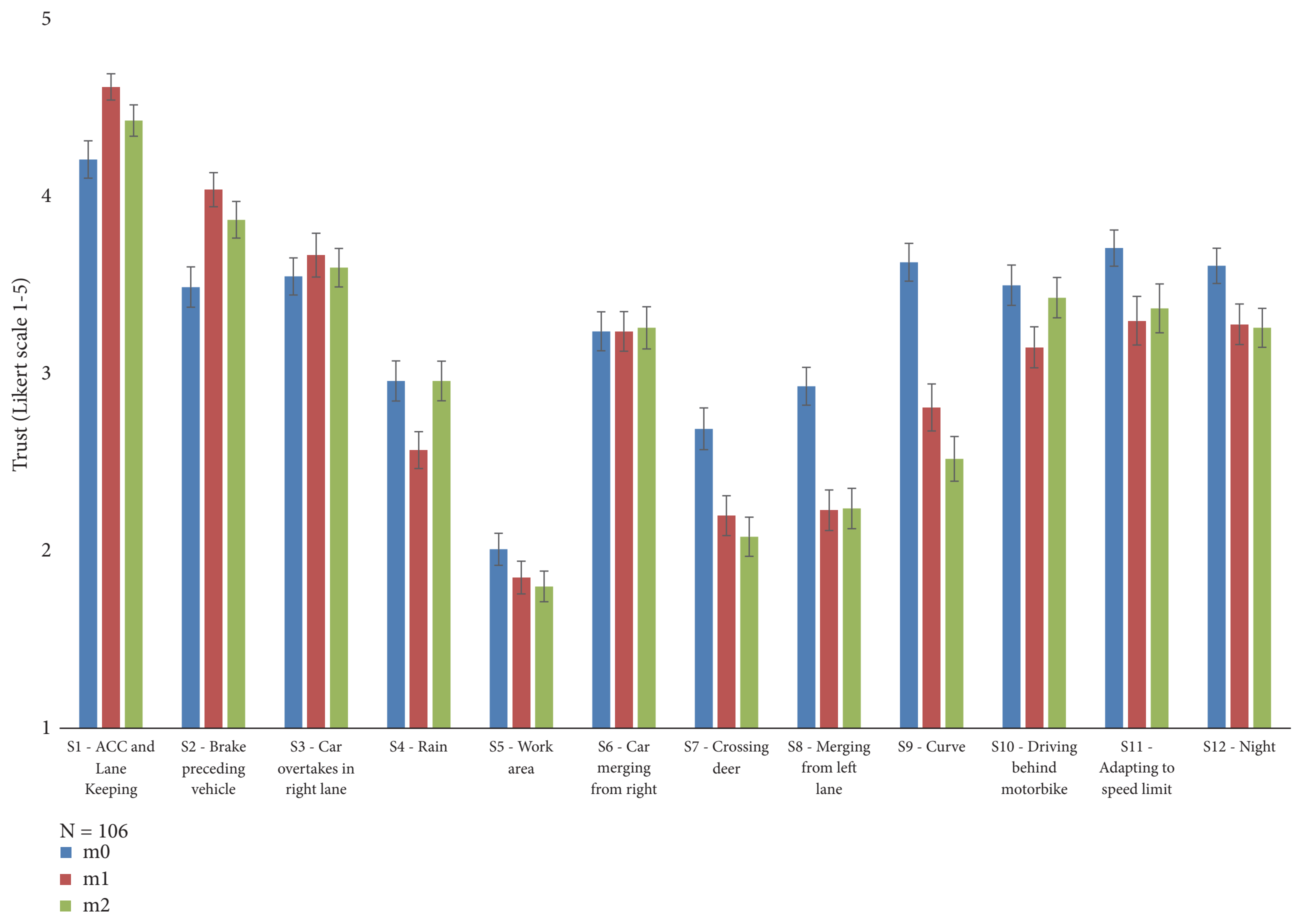

FIGURE 3: **Questionnaire results. Scenarios (S) one to twelve are displayed from left to right, respectively.** For scenarios 1 and 2, drivers' trust significantly increased after experiencing the vehicles. Conversely, for scenarios 4, 7, 8, 9, and 12, drivers' trust significantly decreased after experiencing the vehicles. No significant trust changes between pre- and postmeasurements were found for scenarios 3, 5, 6, and 11. For scenario 10, significant differences were found between the pre- and the "2-week" measurement (i.e., m2), but not between m0 and m1. $\alpha = .004$. Error bars were calculated based on the standard error of the mean.

A Friedman test showed no significant differences across measurements for scenarios 3, 5, 6, and 11 (see Figure 3 and Table 1). The relatively high trust scores observed for scenario 3 (car overtakes in right lane) indicate that drivers trusted the vehicles to cope with this scenario and that this did not change after their experience with the vehicles. A car overtaking in the right lane represents a dangerous situation. Nevertheless, Autopilot, Drive Pilot, and Pilot Assist would be able to detect the oncoming vehicle. Even if drivers did not directly experience scenario 5 (work area), the selected Level 2 vehicles do not perform well with mixed white and yellow road markings.

In scenario 6 (car merging from right), the absence of an effect across measurements suggests that participants are unsure of how the systems can handle the scenario. In scenario 11 (adapting to speed limit), the absence of any significant difference between the three measurements is probably due to the fact that most participants did not have the chance to experience the scenario. In reality, of the vehicles used in the study, only the Mercedes E class was capable of adapting to speed limits.

4. Discussion

While engineering developments are moving fast, there is still little understanding of how drivers interact with automated driving technology [7]. The study of drivers' trust in automation represents a fundamental challenge for current and future human factors research. Importantly, the goal should not be to maximize users' trust toward the technology, but to align drivers' subjective feelings of safety with the actual reliability of the vehicle. In this respect, a decrease in trust is not a negative outcome *per se*: If, for example, vehicles are incapable of handling a specific driving scenario (e.g., deer crossing the road), decreases in trust suggest an improvement in trust calibration. More generally, if drivers overestimate the capabilities of the system, a better understanding of vehicles' limitations after vehicle experience can lead to a decrease in trust. Conversely, if drivers underestimate the capabilities of the system, a better understanding of the vehicles' functionalities can lead to increased trust [11].

Bearing these considerations in mind, determining the "correct" level of trust for particular driving scenarios is a

Table 1: **Questionnaire results**. A Friedman test was performed for an analysis of the variances of m0 (measurement 0), m1 (measurement 1), and m2 (measurement 2). Wilcoxon Signed Ranks test was used for pairwise comparisons (i.e., m0 versus m1, m0 versus m2, and m1 versus m2); *significant ($p <= .004$).

								N = 106
Scenario	**Missing values**	**Friedman test**	**M0 (M, SD)**	**M1 (M, SD)**	**M2 (M, SD)**	**M0 vs m1**	**M0 vs m2**	**M1 vs m2**
(1) You are driving on a highway with an average amount of traffic. The system is set to automatically keep distance from the vehicle in front, and to keep within the lane.	4	$\chi^2 = 23.096$, $p < .001^*$	M = 4.21, SD = 1.09	M = 4.62, SD = 0.76	M = 4.43, SD = 0.92	Z = -3.927, $p < .001^*$	Z = -2.511, $p = .012$	Z = -2.360, $p = .018$
(2) You drive on the highway and the vehicle in front brakes hard.	3	$X^2 = 26.315$, $p < .001^*$	M = 3.49, SD = 1.18	M = 4.04, SD = 0.99	M = 3.87, SD = 1.06	Z = -4.014, $p < .001^*$	Z = -3.390, $p = .001^*$	Z = -2.024, $p = .043$
(3) A car overtakes you in the right lane.	1	$X^2 = 2.537$, $p = .281$	M = 3.55, SD = 1.07	M = 3.67, SD = 1.27	M = 3.60, SD = 1.12	Z = -1.065, $p = .287$	Z = -.444, $p = .657$	Z = -.676, $p = .499$
(4) You drive on the highway with a lot of rain, so the visibility is poor.	0	$X^2 = 6.483$, $p = .039$	M = 2.96, SD = 1.18	M = 2.57, SD = 1.07	M = 2.71, SD = 1.15	Z = -3.390, $p = .001^*$	Z = -2.212, p = .027	Z = -1.484, p = .138
(5) You drive in a work area with yellow road marking in addition to the normal white road marking.	2	$X^2 = 8.130$, $p = .017$	M = 2.01, SD = 0.91	M = 1.85, SD = 0.95	M = 1.80, SD = 0.89	Z = -1.883, p = .06	Z = -2.211, p = .027	Z = -0.668, p = .504
(6) A car merges in from the right lane and has a much lower speed than yours.	1	$X^2 = 0.087$, $p = .957$	M = 3.24, SD = 1.14	M = 3.24, SD = 1.15	M = 3.26, SD = 1.23	Z = -.092, $p = .926$	Z = -.327, p = .744	Z = -.235, $p = .814$
(7) A deer suddenly crosses the highway – a situation in which normally you would brake.	1	$X^2 = 19.196$, $p < .001^*$	M = 2.69, SD = 1.21	M = 2.20, SD = 1.16	M = 2.08, SD = 1.14	Z = -3.437, p = .001*	Z = -4.026, p < .001*	Z = -1.233, p = .218
(8) The lane where you drive ends, you need to insert the lane to your right.	0	$X^2 = 31.848$, $p < .001^*$	M = 2.93, SD = 1.10	M = 2.23, SD = 1.18	M = 2.24, SD = 1.17	Z = -4.658, p < .001*	Z = -4.641, p < .001*	Z = -0.329, p = .742
(9) You are approaching a curve on the highway.	4	$X^2 = 48.771$, $p < .001^*$	M = 3.63, SD = 1.10	M = 2.81, SD = 1.37	M = 2.52, SD = 1.30	Z = -5.239, p < .001*	Z = -6.326, p < .001*	Z = -2.556, p = .011
(10) You are driving on the highway behind a motorbike instead of behind a car.	1	$X^2 = 13.076$, $p = .001^*$	M = 3.5, SD = 1.17	M = 3.43, SD = 1.20	M = 3.15, SD = 1.17	Z = -0.578, p = .564	Z = -3.003, p = .003*	Z = -2.970, p = .003*
(11) You are driving on the highway and the speed limit decreases. Your speed should adapt to the new speed limit.	3	$X^2 = 2.299$, $p = .317$	M = 3.71, SD = 1.06	M = 3.30, SD = 1.41	M = 3.37, SD = 1.42	Z = -2.402, $p = .016$	Z = -2.253, $p = .024$	Z = -.405, $p = .686$
(12) You are driving at night on the highway.	0	$X^2 = 13.292$, $p = .001^*$	M = 3.61, SD = 1.02	M = 3.28, SD = 1.18	M = 3.26, SD = 1.13	Z = -2.989, p = .003*	Z = -3.214, $p = .001^*$	Z = -0.180, p = .857

challenging task, particularly in realistic driving conditions where road conditions change continuously and unexpected issues can compromise the safety of the driver at any moment. Nonetheless, observed changes in drivers' trust after vehicle experience were in the direction we would expect if drivers had improved their understanding of vehicles' limitations and capabilities.

More specifically, the results of our study suggest an improvement in trust calibration in 7/12 scenarios. In two scenarios (i.e., car overtakes in right lane; work area) drivers already showed correct trust calibration and this was maintained after the driving experience. In four scenarios, including the two in which trust was already well calibrated, the changes observed were not significant.

Notably, improved trust calibration involved both increases and decreases in trust. This is a key finding, showing that on-road driving experience positively influenced drivers' trust by increasing participants' ability to discriminate between different driving situations. Importantly, this was also true for scenarios (e.g., crossing deer) that were not directly experienced by any of the participants. Thus, vehicle experience leads to more reliable inferences concerning both experienced and unexperienced scenarios.

More in detail, after vehicle experience, drivers showed *increased trust* in vehicles' ACC and Lane Keeping capabilities, which are normally considered to be reliable. In the other five scenarios where we observed a change in trust, we observed a *reduction in trust*, again reflecting a better understanding of the vehicles' capabilities. Specifically, drivers seem to have understood that weather conditions can have a major impact on the reliability of the system, that they cannot rely on the emergency braking system when a deer suddenly crosses the highway, that current Level 2 vehicles do not have the capability to merge autonomously, that partially automated vehicles with Lane Keeping functionality can stray out of lane when facing tight curves, and that drivers should not blindly rely on automation in low visibility conditions.

When asked if the system could work optimally when driving behind a motorcycle (instead of behind a car), drivers' trust did not change immediately after experiencing the vehicles. However, in the "2-week" measurement, drivers' reported trust was significantly lower. In a recent study by Lenkeit [21], motorcycles were inadequately detected in more than 40% of cases (against a ~4% failure rate for detection of cars). In sum, vehicle detection systems are still far from perfect. Even if this scenario was not directly experienced by our participants, drivers' trust was better aligned to the actual reliability of the systems after their on-road driving experience. The fact that participants' decrease in trust occurred only two weeks after experiencing the vehicles could indicate that they needed more time to process this specific scenario. Nevertheless, since trust did not change immediately after vehicle experience, we cannot decisively conclude that the driving session was the main factor leading to improved trust calibration.

Participants' initial trust scores (i.e., before their driving experience) suggest a general overestimation of vehicles' capabilities. Drivers had no experience with the selected systems and their expectations of system behaviour were probably influenced by the media. "Self-driving cars" are currently a hot topic, leading to increased awareness of the presence and promises of the technology [10, 22]. News agencies, dealers, and manufacturers often suggest that fully autonomous vehicles will soon be cruising our roads (e.g., Tesla's website opens with the statement "Full Self-Driving Hardware on All Cars"). In reality, many analysts believe that full automation is still decades away (e.g., [23–25]). Thus, the initial overestimation of vehicles' capabilities seen in our participants might be the result of erroneous beliefs concerning the status of available driving technology.

In the scenarios where we found differences between the pre- and postmeasurement, no differences were found between the post- and the "2-week" measurement (see Table 1). This suggests that the impact of the driving experience on drivers' trust was relatively stable. No differences between the pre- and the "2-week" measurement were found when drivers were questioned about the ACC and Lane Keeping functions of the Level 2 vehicle (i.e., scenario 1), or when they were asked if the system could function optimally in low visibility conditions due to heavy rain (i.e., scenario 4). In these two scenarios, the driving experience had a direct positive impact on drivers' trust, but the effect did not seem to last. In the other scenarios, where differences between the pre- and postmeasurements were found, we also found differences between the pre- and the "2-week" measurement.

Our study presents several limitations that should be acknowledged. Participants drove two different vehicles before completing the post questionnaires (i.e., immediately after the driving session and two weeks after the driving session). These vehicles were all categorized as Level 2 and presented comparable Level 2 functions. Nevertheless, we cannot assess from the questionnaire data if one vehicle influenced trust scores more than the other. The event started with a presentation, in which the concept of partial automation was introduced to the participants for the first time. Furthermore, during every drive, a vehicle expert explained the Level 2 functions available in the vehicle and prompted the drivers to use them. We cannot assess if trust scores were mainly influenced by the participants' on-hand driving experience, by the initial presentation, by the experts' feedback, or by a combination of these factors. Finally, data was collected during five separate driving sessions. Varying weather conditions between sessions may have affected drivers' trust scores.

To our knowledge, this study represents the first attempt to measure changes in trust after drivers' experience with real Level 2 cars. Overall, our results suggest that hands-on experience with Level 2 vehicles helps trust calibration. Drivers interested in purchasing a Level 2 vehicle should be given the chance to drive the car along with someone (e.g., car dealer) who is fully aware of its functionalities. Such a procedure is likely to improve driver's understanding of the vehicles' capabilities and contribute to trust calibration even in situations they did not encounter during their on-road experience.

5. Conclusions

In this study we analysed, for the first time, participants' trust before (i.e., premeasurement), immediately after (i.e., postmeasurement), and two weeks after (i.e., "2-week" measurement) on-road driving experience with SAE Level 2 semiautomated vehicles. The results showed that the real-life driving experience led to improvements in trust calibration. Drivers' ability to more precisely differentiate, after experience, between situations where the automation performed well from situations where it did not, played a key role in calibrating participants' trust. These improvements were experienced by all drivers, independent of age, gender, and innovation profile. We conclude that carefully organised instruction and on-road experience can make crucial contributions to safe driving, while simultaneously improving sales, encouraging uptake, and helping to realize the potential benefits of self-driving technology for drivers and society as a whole.

Data Availability

The data collected during the study are freely available at osf.io/f7rm9.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This research was completely funded by Rijkswaterstaat, The Ministry of Infrastructure and Water Management, the Dutch road authority. The authors thank Rijkswaterstaat for providing the opportunity to conduct this research.

References

[1] A. Tigadi, R. Gujanatti, and A. Gonchi, "Advanced driver assistance systems," *International Journal of Engineering Research and General Science*, vol. 4, no. 3, pp. 151–158, 2016.

[2] K. A. Brookhuis, D. De Waard, and W. H. Janssen, "Behavioural impacts of advanced driver assistance systems – an overview," *European Journal of Transport and Infrastructure Research*, vol. 1, no. 3, pp. 245–253, 2001.

[3] O. Gietelink, J. Ploeg, B. de Schutter, and M. Verhaegen, "Development of advanced driver assistance systems with vehicle hardware-in-the-loop simulations," *Vehicle System Dynamics*, vol. 44, no. 7, pp. 569–590, 2006.

[4] J. C. F. de Winter, R. Happee, M. H. Martens, and N. A. Stanton, "Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 27, pp. 196–217, 2014.

[5] S. A. E. Taxonomy, "Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems," Tech. Rep., SAE International, 2014.

[6] C. Gold, M. Körber, C. Hohenberger, D. Lechner, and K. Bengler, "Trust in automation–before and after the experience of take-over scenarios in a highly automated vehicle," *Procedia Manufacturing*, vol. 3, pp. 3025–3032, 2015.

[7] M. H. Martens and A. P. van den Beukel, "The road to automated driving: dual mode and human factors considerations," in *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems: Intelligent Transportation Systems for All Modes (ITSC '13)*, pp. 2262–2267, IEEE, October 2013.

[8] J. K. Choi and Y. G. Ji, "Investigating the Importance of Trust on Adopting an Autonomous Vehicle," *International Journal of Human-Computer Interaction*, vol. 31, no. 10, pp. 692–702, 2015.

[9] M. Ghazizadeh, J. D. Lee, and L. N. Boyle, "Extending the Technology Acceptance Model to assess automation," *Cognition, Technology & Work*, vol. 14, no. 1, pp. 39–49, 2012.

[10] N. Trübswetter and K. Bengler, "Why should I use ADAS? Advanced driver assistance systems and the elderly: knowledge, experience and usage barriers," in *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, pp. 17–20, Bolton Landing, NY, USA, June 2013.

[11] B. M. Muir, "Trust between humans and machines, and the design of decision aids," *International Journal of Man-Machine Studies*, vol. 27, no. 5-6, pp. 527–539, 1987.

[12] E. J. de Visser, M. Cohen, A. Freedy, and R. Parasuraman, "A design methodology for trust cue calibration in cognitive agents," in *International Conference on Virtual, Augmented and Mixed Reality*, pp. 251–262, Springer, Cham, Switzerland, 2014.

[13] J. D. Lee and K. A. See, "Trust in automation: designing for appropriate reliance," *Human Factors*, vol. 46, no. 1, pp. 50–80, 2004.

[14] M. S. Cohen, R. Parasuraman, D. Serfaty, and R. Andes, *Trust in Decision Aids: A Model and a Training Strategy*, Cognitive Technologies Inc, Arlington, Va, USA, 1997.

[15] J. Lee and N. Moray, "Trust, control strategies and allocation of function in human-machine systems," *Ergonomics*, vol. 35, no. 10, pp. 1243–1270, 1992.

[16] B. Rajaonah, F. Anceaux, and F. Vienne, "Trust and the use of adaptive cruise control: A study of a cut-in situation," *Cognition, Technology & Work*, vol. 8, no. 2, pp. 146–155, 2006.

[17] D. A. Dickie and L. N. Boyle, "Drivers' understanding of adaptive cruise control limitations," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 53, no. 23, pp. 1806–1810, 2009.

[18] A. Koustanaï, V. Cavallo, P. Delhomme, and A. Mas, "Simulator training with a forward collision warning system: effects on driver-system interactions and driver trust," *Human Factors*, vol. 54, no. 5, pp. 709–721, 2012.

[19] M. Körber, E. Baseler, and K. Bengler, "Introduction matters: Manipulating trust in automation and reliance in automated driving," *Applied Ergonomics*, vol. 66, pp. 18–31, 2018.

[20] B. M. Muir, "Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems," *Ergonomics*, vol. 37, no. 11, pp. 1905–1922, 1994.

[21] J. F. Lenkeit, "Preliminary study of the response of forward collision warning systems to motorcycles," in *Proceedings of the 11th International Motorcycles Conference*, Cologne, Germany, October 2016.

[22] V. Hahanov, W. Gharibi, K. L. Man et al., "Cyber-physical technologies: hype cycle," in *Cyber Physical Computing for IoT-driven Services*, pp. 259–272, Springer, Cham, Switzerland, 2018.

[23] P. Bansal and K. M. Kockelman, "Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies," *Transportation Research Part A: Policy and Practice*, vol. 95, pp. 49–63, 2017.

[24] A. Lumiaho and F. Malin, "Road transport automation: road map and action plan 2016-2020," Research Reports, Finnish Transport Agency, 2016.

[25] T. Litman, *Autonomous vehicle implementation predictions*, Victoria Transport Policy Institute, 2018.