Trust in an Autonomously Driven Simulator and Vehicle Performing Maneuvers at a T-Junction with and Without Other Vehicles

Phillip L. Morgan[1](✉), Craig Williams[2], Jonathan Flower[2], Chris Alford[2], and John Parkin[2]

[1] School of Psychology, Cardiff University,
70 Park Place, Cardiff CF10 3AT, UK
morganphil@cardiff.ac.uk
[2] University of the West of England – Bristol,
Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK
{craig6.williams,jonathan.flower,chis.alford,
john.parkin}@uwe.ac.uk

Abstract. Autonomous vehicle (AV) technology is developing rapidly. Level 3 automation assumes the user might need to respond to requests to retake control. Levels 4 (high automation) and 5 (full automation) do not require human monitoring of the driving task or systems [1]: the AV handles driving functions and makes decisions based on continuously updated information. A gradual switch in the role of the human within the vehicle from active controller to passive passenger comes with uncertainty in terms of trust, which will likely be a key barrier to acceptability, adoption and continued use [2]. Few studies have investigated trust in AVs and these have tended to use driving simulators with Level 3 automation [3, 4]. The current study used both a driving simulator and autonomous road vehicle. Both were operating at Level 3 autonomy although did not require intervention from the user; much like Level 4 systems. Forty-six participants completed road circuits (UK-based) with both platforms. Trust was measured immediately after different types of turns at a priority T-junction, increasing in complexity: e.g., driving left or right out of a T-junction; turning right into a T-junction; presence of oncoming/crossing vehicles. Trust was high across platforms: higher in the simulator for some events and higher in the road AV for others. Generally, and often irrespective of platform, trust was higher for turns involving an oncoming/crossing vehicle(s) than without traffic, possibly because the turn felt more controlled as the simulator and road AVs always yielded, resulting in a delayed maneuver. We also found multiple positive relationships between trust in automation and technology, and trust ratings for most T-junction turn events across platforms. The assessment of trust was successful and the novel findings are important to those designing, developing and testing AVs with users in mind. Undertaking a trial of this scale is complex and caution should be exercised about over-generalizing the findings.

Keywords: Autonomous driving · Autonomous simulator
Autonomous vehicle · Trust · Acceptability

1 Introduction

The technological development of road vehicles that can drive themselves is accelerating at a swift pace [5, 6], much like former boom periods in other domains such as aviation and nuclear. Within these domains, some automated systems were not always developed with the human operator/monitor/user in mind, and this led to multiple problems [7–10]; including abuse, overuse, misuse, and in some cases, disuse [11]. Automation disuse is often associated with lack of trust in the system to perform the job (s) it was designed for, with trust influencing acceptability, adoption and continued use of automation technology [2, 12]. Trust in automation is influenced by factors such as the reliability, resilience, and robustness of the system [13, 14]. Multiple frameworks have been developed to better understand automation in terms of appropriate levels and types to support optimal human interaction and most consider the key issue of trust [11, 14–17]. The widespread use of road vehicles that can drive themselves is rapidly approaching: many with possible fallback to human users [1] and some under development that are highly (Level 4) or even fully autonomous (Level 5) and do not require human intervention or even monitoring of the driving task or systems. Despite this, few AV studies over the past decade or so have measured trust [3, 4]. Given past lessons learned, trust should feature as a major factor when testing and developing AVs whether using simulator or actual road-vehicle platforms. The main aim of the current study is to measure human trust in simulator and road-based AVs that are able to perform a series of increasingly complex maneuvers, with or without other traffic.

Trust is quite easily one of the most important enablers (and indeed barriers) to humans adopting and continuing to use new automation technology. It is a key parameter of the Automation Acceptance Model [15] that considers trust in automation in terms of intention to use [18], adoption [19, 20], reliance [21, 22], and possible rejection due to an untrustworthy experience or experiences [23]. These studies stress the multi-dimensionality of trust as a construct based upon experiential factors such as system reliability, predictability, ability to efficiently handle all associated actions, as well as individual characteristics such as propensity to trust (including trust in technology and in automation, [15]). As important as it is to consider, trust is subject to difficulties in terms of how and when to measure, as well as human individual differences.

Some recent studies have considered human trust when testing mainly Level 3 AV technology in the laboratory [3, 4, 24] and on the road [25]. [3] conducted a simulator study in AV mode with a maximum speed of 20-mph and measured trust and comfort when passing other objects including bicycles and scooters. They found that trust and comfort were highest during earlier steering maneuvers and with wider lateral distances than participants reportedly would implement themselves during manual driving. Despite these interesting findings, the authors did not test participants in a control simulated manual non-AV driving mode, and, trust and comfort ratings were self-reported and thus possibly impacted by idiosyncratic subjective factors.

Some studies have examined trust in Level 3 simulation AV systems that handback control of the AV system at various points within a journey, with subsequent switching back again to autonomous mode. [6] define takeover as the time taken to re-engage

with vehicle controls and handover as the time taken to regain a baseline/normal level of driving. [4] conducted a multi-phase simulator study and measured trust in automation before as well as after experiencing handover scenarios. They found a general trend that experience led to an increase in self-reported trust in automation, although this was marginally non-significant. They also examined age as a possible mediating factor (given that age can influence trust in automation: [16]) and found a significant improvement in trust for participants aged above 60 years which was not significant for those 30 years or younger. They attributed this difference to older adults relying more on automation [26] and tending to be a population sector most sensitive to automation reliability changes [27]. The scenarios used were however limited to high speed (120 km/hr) freeway type driving for 15–20-min periods with relatively infrequent handovers. In fact, the Level 3 handover design (i.e., request occurs x seconds in advance of a potential collision) could have affected findings as even though handover requests were unpredictable, participants would have likely learned to expect them thus being poised to retake driving controls. Thus, studies that do not include handover are needed. Furthermore, the authors measured trust after completion of each scenario rather than during scenarios with the latter likely to capture more accurate situation-specific ratings. However, capture immediately after the event may distract participants but many real-time measures (e.g., situation awareness/SA [28, 29]) have similar issues.

A recent study by [24], again using a handover paradigm but with non-critical as well as critical takeover situations, investigated the impact of introductory instructions designed to increase ('trust promoted') or decrease ('trust lowered') on reliance of and trust in Level 3 vehicle autonomy. They found higher ratings of trust with experience although only moderate differences due to the manipulation of introductory information. For example, those in the trust promoted group spent more time looking at a non-driving related task, and were more likely to over-rule the AV system in non-critical situations. Alarmingly, they also found that a sub-set of trust promoted participants collided with obstacles compared to none in the trust lowered group.

Finally, it is worth noting that AV trust has recently been considered outside of the laboratory. [25] reported on a naturalistic study using the *Tesla Model S* over 6-months. At the end of the study, a cautionary note was stressed in terms of over-trust in vehicles with self-driving capabilities. Specifically, and drawing upon her own work on SA spanning more than two decades [30–32], noted that increased reliability will lead to increased trust. This will likely have a damaging effect on SA (e.g., comprehend current situation, project future states), which could be problematic for Level 3 AVs at least in the event of e.g., a handover request. This unearths a crucial dilemma in which it might actually be counterintuitive to strive for high levels of trust in Level 3 AVs although this should be highly desirable in Level 5 (and possibly Level 4) AVs.

Current Study. The main aim of the current experiment was to measure human trust in AVs – one road based, one simulator – performing a series of frequently experienced T-junction maneuvers (with and without approaching/oncoming vehicles). In non-autonomous vehicles, this requires fine-tuned driving skills and experience and a series of complex cognitive processes including perception, attention, memory and judgment. Maneuvering at a T-junction is a very common urban road situation. The AV needs to

safely and efficiently handle each one of the different possible movements, both with and without on-coming traffic. In particular, the vehicle's ability to decide whether it is safe to make a turn or not is critical. We tested a variety of T-junction turn events with increasing complexity from empty roads to turns involving another vehicle(s). For this, and to our knowledge, first study of its type, the AVs always yielded to oncoming/crossing vehicles before making a turn into or out of a side road at a T-junction.

There are a number of predictions. First, trust ratings will be higher for all events in the simulator compared with the road-based AV owing to the fact that the former is a fixed non-moving platform. Second, turns associated with the highest degree of perceived risk (i.e., turning into or out of a side road with at least one oncoming/crossing vehicle) will result in the lowest trust ratings, especially within the road-based AV where participants may perceive the chances of a road traffic collision to be higher. Related to this is that the prediction that the lowest trust ratings should be associated with turns involving more than one oncoming/crossing vehicle. Nevertheless, and given the novelty of the current experiment, and noting that participants had no past direct experience of an AV, these predictions are tentative at best. We included questionnaires that measure trust in technology and automation. It is predicted that higher ratings on such questionnaires will be related to higher trust ratings for simulator and road-based T-junction maneuvers; especially between trust in automation and trust ratings within the road-based AV. Despite this, we also hypothesize that such positive relationships will be weaker as the complexity associated with the T-junction maneuver increases.

2 Method

Participants. A quota sampling method was used to recruit 46 volunteers aged 22–78 years of age (M = 46.22; SD = 15.53). Twenty were women. The sample size was adequate to detect medium-large effect sizes (Cohen's f = .25 −.4) with power of .8 [33]. All had full driving licenses and experience of driving in the UK, ranging from 2–60 years (M 26.26; SD 18.56). All had normal-corrected vision and hearing, and were English first language or highly proficient in English as a second language. The highly immersive simulator resulted in some experiencing nausea and two had simulator sickness. The simulator drop-out rate after one circuit was 11% rising to 24% after three circuits. Thus, a reduced (N = 37, adequate to detect large effect sizes) sample were considered for cross-platform analyses. None experienced high levels of nausea or sickness within the road AV, although one did not complete due to not feeling comfortable.

Design. A repeated measures design was adopted whereby participants were driven on three circuits of the same ∼10-min route and thus experienced the same events (Table 1) in the same order of increasing complexity (route permitting: Event 3, 5, 4, 6, 1, 7, 2) three times. There was also a shorter practice route to provide orientation with the simulator and road AV. There were seven Event Types of T-junction maneuvers per circuit (see Table 1) and thus 21 T-junction maneuvers in total for participants who

completed all three circuits, 14 for those who completed two full circuits, and so on. The main dependent measure was trust rating in the simulator and road AV recorded immediately after completion of each T-junction maneuver. This was measured on an 11-point Likert scale ranging from 0 (no trust) to 10 (complete trust). Nausea ratings were taken immediately after each circuit (including practice), again using an 11-point Likert scale ranging from 0 (no nausea) to 10 (completely nausea). Platform order was counterbalanced such that 50% of participants completed the simulator component first.

Materials. The hardware consisted of a *Williams Advanced Engineering* modified *Land Rover Evoque Sport* fixed base simulator and a bespoke *Land Rover Bowler* 'Wildcat' autonomous road vehicle (Fig. 1). The Wildcat contains actuators and additional braking system and is governed by multiple e-stops configured to either stop the vehicle and apply brakes or revert to manual control. The Wildcat is programmed using real-time GPS and sensor data to follow and learn pre-planned routes and within the bounds of the road layout until it is able to perform optimally in full AV mode. All decisions (e.g., when to slow down and at what rate, when to pull in or out of junctions and at what speed, level of assertiveness) are controlled by a bespoke Decision Making System (DMS) designed and programmed by partners working on the current project (Bristol Robotics Laboratory and BAe Systems). The DMS used: the finite state machine (allowing it to know what part of the circuit it was on); a clear distance measurement (to stop behind the obstacle car for an avoidance maneuver); and, vehicle crossing detection (to give a clear-to-go signal at junctions). This scripted approach meant that decisions were constrained and consistent. The critical gap acceptance at T-junctions was set to 4 s which is the time accepted by 50% of drivers [34], and both the simulator and Wildcat were programmed to yield to approaching vehicles at all times.

Fig. 1. Wildcat road AV (left) and Williams advanced engineering AV simulator (right)

Highly immersive simulator journeys were programmed to mimic Wildcat scenarios using *Oktal (Simulation in Motion)* software modified by BRL programmers. The simulator used the same autonomous DMS as the Wildcat. The set-up included three large projector screens to provide 180º front and side views, and side-mirrors with back left and right screens projected, and a windscreen mounted rearview mirror with the rear view projected via a large monitor. The interior was standard for the vehicle model. The simulator was controlled by five *Hewlett Packard* 8 Core 3.70 GHz *Intel*

Xeon v3 PCs. There was also an experimenter control station with 5 21" *Iiyama Prolite E2480HS* monitors. The DMS ran on a separate PC of the same specification.

A ~10-min driving circuit involving four different routes (loops) around a major carpark (including approach and exit roads) within the University of the West of England, Bristol (UK) Frenchay campus was designed. The vehicle and simulator stopped after each circuit (not between loops) and started the next circuit when the participant was ready to continue. Events occurred (Wildcat) or were programmed to occur (simulator) within the same positions during each circuit and included seven different instances of negotiating a set of T-junctions (Table 1). Event 1 involved turning right off the main road into the side road and differed to Event 2 which involved yielding to an on-coming vehicle before making a right turn. Event 7 was similar to Event 2 but involved turning left into the side road and did not need a yield to be performed by the vehicle. Event 3 involved turning left out of the side road onto the main road, and Event 4 was similar although involved the AV yielding before the turn to allow a vehicle to pass on the main road. Event 5 was similar to Event 3 although involved a right turn and Event 6 was similar to Event 5 but involved yielding before committing to the turn due to vehicles passing in each direction on the main road. This pattern of events allowed for three direct comparisons: Events 1 and 2 at the same right hand turning; Events 3 and 4 at the same left hand turn out; and, Events 5 and 6 at the same right hand turn out. There were variable lengths of time interval between events, which helped reduce anticipation and possible unintentional fluctuations in alertness.

Table 1. Scenarios and events

Event Number	Event Description	Event Illustration
1	Turning right off the main road into the side road at a priority junction with no other vehicles.	
2	Turning right off the main road into the side road at a priority junction with an on-coming vehicle approaching on the main road.	
3	Turning left out of a side road onto a main road at a priority junction with no other vehicles.	
4	Turning left out of a side road onto a main road at a priority junction with a vehicle approaching on the main road from the right.	
5	Turning right out of a side road onto a main road at a priority junction with no other vehicles.	
6	Turning right out of a side road onto a main road at a priority junction with vehicles approaching along the main road from both directions.	
7	Turning in left into a side road from a main road at a priority junction with no	

The Wildcat and simulator were programmed to drive with a 'neutral' (not assertive or cautious) driving personality, that is with a gap acceptance at junctions of 4-s. Similarly, vehicle acceleration, including starting from a standstill, was programmed to occur in a neutral non-assertive manner. Both platforms were set to maximum speeds of 20-mph and only achieved this on longer straight or slightly curved road stretches, which is the campus speed limit and consistent with many UK urban city center roads.

Some questionnaires and scales were administered. Key to the current study are the:

– General Trust (in Technology) Scale/GTS [35], which contains seven questions (e.g., 'I believe that most technologies are effective at what they are designed to do') with 7-point Likert scale answers. A higher overall score represents higher trust in technology.
– Trust in Automation Checklist/TAC [36], which contains 12 questions (e.g., 'the system is reliable') measuring trust in the autonomous platform just experienced with higher scores indicating increased dependability and trust in the system.

Procedure. Participants were given a detailed pre-experiment briefing including health and safety prior to consenting to take part. For the simulator component, participants sat in the right-hand driving seat and could adjust the seat position. A researcher sat in the passenger seat to record trust and nausea ratings. A 2.5-min practice and orientation loop were experienced followed by a nausea rating request. A nausea rating of ≥ 1–4 resulted in asking the participant if they were comfortable continuing, and advice was given not to continue if a rating of ≥ 5 was given and/or if the participant felt sick. For those that could continue (>98%) circuit 1 began, and the experimenter called out the trust rating question immediately after each event had been experienced, and recorded the participant's rating. The full question was: "On a scale of 0–10, where 0 is 'no trust' and 10 is 'complete trust', rate how much you trusted the automated vehicle simulator during the last maneuver." As participants became familiar with the procedure, this was simplified until only needing to ask "rate trust". Another nausea rating was taken at the end of the first circuit. There were 1–2-min breaks between circuits, and the same testing protocol applied for circuits 2 and 3. Participants could ask to stop the experiment at any time and the simulator could be stopped immediately either by the experimenter sitting with them and/or by a second experimenter sitting at the control station. This component of the experiment took approximately 40–45 min.

The Wildcat procedure was very similar. It involved an additional safety briefing to familiarize participants with a safety driver who could take manual control at any point and/or stop the vehicle in the event of an emergency or emergency stop protocol, which could be activated at any time during a journey by the participant. In order to meet the requirements of the safety case and maximize safe operation, the Wildcat component involved deployment of marshals around the circuit who were in constant contact with 2–3 experimenters in a control center with visibility of the test track and the safety driver. The chief control center experimenter requested all trust and nausea ratings via an audio communication system linked to headphones worn by the participant. Verbal trust and nausea ratings were recorded via a microphone and logged by control center experimenters. This component of the experiment took approximately 50–55 min. At the end of the experiment, all participants received a full written and verbal debrief.

3   Results and Discussion

Figure 2 displays mean trust ratings (over two circuits) for each of the events experienced across both platforms. Generally, trust ratings were quite high (lowest = 7.39/10 for simulator Events 1 and 5, highest = 8.37/10 for simulator Event 4), and marginally highest overall for the simulator platform ($M = 7.86$ versus 7.76 for the Wildcat).
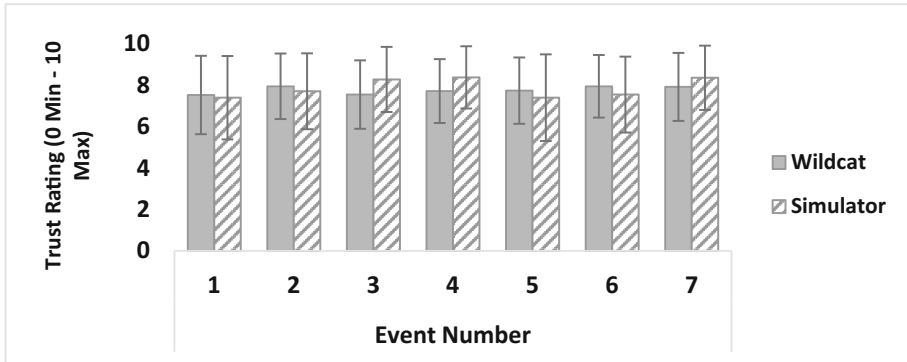

Fig. 2.   Mean trust rating (Two Circuits) by event type. Error bars = ±Standard deviation.

For the Wildcat platform, trust ratings seem surprisingly higher for Events 2 and 6 (right turning with an oncoming vehicle(s)) versus Events 1 and 5 (right turning without an oncoming vehicle(s)). This was not the case for left turn events where trust ratings were similar with and without other traffic (Events 3, 4 and 7). For the simulator platform, trust is (perhaps surprisingly) higher for Events 2 and 6 (right turn with oncoming vehicle(s)) versus Event 1 (right turn without an oncoming vehicle(s)), although the rating is similar for Events 5 and 6 (right turn with and without an oncoming vehicle(s)). Despite higher trust ratings for left versus right turns, there appears to be little difference between Event 4 (that involved oncoming traffic) and Events 3 and 7 (where the route was clear). Variance (as measured using standard deviations) is relatively low across all events and between platforms, although perhaps surprisingly highest for Event 5 in the simulator (turning right without an oncoming vehicle(s)).

Within platform trust rating data was analyzed using 2-tailed paired-samples t-tests. For the Wildcat, ratings were significantly higher for Event 2 (turning right with one oncoming vehicle) versus Event 1 (turning right with no oncoming traffic, $t(44) = 3.39$, $p < .001$, and did not support our prediction of an effect in the opposite direction. There was a non-significant difference between Events 3 (turning left without an oncoming vehicle) and 4 (turning left with one oncoming vehicle), $t(44) = 1.40$, $p = .17$, which did not support our prediction of a difference (higher trust for Event 3). Ratings were significantly higher for Event 6 (right turn with crossing traffic) versus Event 5 (right turn with no traffic), $t(44) = 2.68$, $p = .01$, again in the opposite direction predicted.

For the simulator, trust was also significantly higher for Event 2 than 1, $t(37) = 2.72$, $p = .01$, and there was no difference between Events 3 and 4, t(37) = 1.40, $p = .17$. Unlike the Wildcat, there was also no difference between Events 5 and 6, t(37) = 1.57, $p = .12$.

A series of factorial repeated measures analyses of variance (ANOVA) were conducted with platform as one variable (Wildcat versus simulator) and highly similar and comparable Event pairings (i.e., 1 versus 2, 3 versus 4, 5 versus 6) as the other variable within each analysis. Bonferroni tests were applied for all post-hocs. The reduced sample of 37 participants who completed at least the first two circuits within the simulator were included within these analyses. For Event 1 (turning right with no on-coming traffic) versus 2 (turning right with one on-coming vehicle), there was a non-significant main effect of platform, $F(1, 36) = .16$, $MSE = .947$, $p = .70$, a significant main effect of Event, $F(1, 36) = 15.46$, $MSE = .27$, p < .001, $\eta p2 = .30$, and a non-significant interaction, $F(1, 36) = .037$, $MSE = .29$, $p = .73$. Trust ratings were higher for Event 2 than 1 (p < .001), irrespective of platform type, whilst we predicted that they would be higher in the simulator. For Event 3 (turning left without an oncoming vehicle) versus Event 4 (turning left with one oncoming vehicle), there was a significant main effect of platform, $F(1, 36) = 19.36$, $MSE = 1.14$, $p < .001$, $\eta p2 = .350$, due to higher trust within the simulator. This was in line with our prediction. There was a non-significant main effect of Event, $F(1, 36) = 3.05$, $MSE = .21$, $p = .090$, and a non-significant interaction, $F(1, 36) = .584$, $MSE = .20$, $p = .45$. For Event 5 (right turn with no traffic) versus Event 6 (right turn with crossing traffic), there was a non-significant main effect of platform, $F(1, 36) = 2.06$, $MSE = .91$, $p = .159$, despite our prediction that trust would be higher in the simulator. There was a significant main effect of Event, $F(1, 36) = 4.82$, $MSE = .17$, $p = .04$, $\eta p2 = .11$ as trust ratings were higher for Event 6 than 5, irrespective of platform type. The interaction was not significant, $F(1, 36) = .04$, $MSE = .16$, $p = .84$.

Finally, we consider possible correlations (Pearson's r) between participant factors (age, driving experience, driving time over past year, trust in technology, and trust in automation) and trust ratings. In terms of participant age, 13 of 14 correlations run (seven events per platform) were non-significant. The only significant correlation in terms of age was for Event 3 in the simulator (turning left without an oncoming vehicle), and this was positive ($r = .31$, $p = .03$) suggesting higher trust ratings for older participants. Next, we considered number of years as a qualified driver and again found only one significant correlation again for Event 3 in the simulator ($r = .34$, $p = .02$). There were no significant correlations for time spent driving over the past year. Trust ratings for most events across both platforms (especially Wildcat) were positively related to trust in technology (Table 2), apart from Event 3 (turning left without an oncoming vehicle) and 6 (right turn with crossing traffic) in the simulator. Also, trust ratings for most Events across both platforms (all Wildcat Events) were positively related to trust in automation (Table 2), apart from Event 3 and 4 (turning left with one oncoming vehicle) in the simulator. Generally, the findings from the latter two sets of correlations were largely expected and reasonable, although a possible limitation in relation to our tested sample is discussed within the limitations section below.

Table 2. Correlations (using Pearson's r Tests) between trust in technology and trust in automation for each event (1–7) across each platform (Simulator and Wildcat)

	Event						
	1	2	3	4	5	6	7
	Trust in technology						
Simulator	.54**	.49**	.28	.33*	.48**	.30	.40**
Wildcat	.44**	.41**	.37**	.42**	.40**	.35**	.50**
	Trust in automation						
Simulator	.37*	.42**	.29	.28	.32*	.35*	.31*
Wildcat	.50**	.42**	.46**	.46**	.50**	.40**	.40**

Note. $* p < .05$, $** p < .01$, no $* p > .05$ (non-significant)

4   Limitations

There are limitations to the current study, many of which were unavoidable consequences of embarking on such a large-scale experiment using both a road-based AV and highly immersive simulator. The scale of the experiment, and requisite levels of sophistication and reliability with both platforms, as well as factors that could not be controlled for, are likely to have impacted upon at least some of the trust ratings. Wildcat factors include change in weather conditions that occasionally meant pausing between circuits, and battery capacity when testing over prolonged periods. However, we are confident that these issues did not have a major impact as standard deviations were relatively low and mostly consistent across platform and Event means, and this was helped by averaging over no less than two full circuits. A further factor with the simulator was the lack of smoothness ('jerkiness') at some junctions which may have impacted results. There was also an issue with simulator nausea and sickness. Only 35 out of 46 participants completed three circuits in the simulator, with 38 completing two, and 41 completing one. Since running the current experiment, the simulator and scenarios have undergone work to reduce both issues and we are developing a better simulator orientation protocol within another project (http://www.flourishmobility. com/) that is proving to be effective. Also, and due to our recruitment method, participants were mostly self-selecting having responded to an advert to take part in an early project-related study regarding views on, attitudes, and expectations of AVs. Thus, many may have been sympathetic to new and emerging technologies, which might have affected trust ratings. Finally, the current reported findings only include one type of frequently experienced vehicle maneuver (T-junctions) and others require consideration (such as negotiating parked vehicles, pedestrians, and other road users such as cyclists).

5   Implications

There are many important implications of the study although we can only cover some in interest of brevity. We have successfully demonstrated how trust can be measured within a highly immersive simulator designed and programmed to mimic a road-based AV. There were multiple instances where trust for T-junction events did not differ between platforms and an almost equal number of instances where it was higher for one platform. Despite issues with simulator nausea and sickness, this is an important step towards validating simulator platforms to orient humans to AVs in terms of increasing experience and trust. This is important based on limits of road-based testing given e.g., infrastructure, costs, and safety. A future step towards developing and testing Level 4 and 5 AVs will be to assess whether even more immersive solutions such as virtual and augmented reality environments could be used. The methods used in the current study represent an important step to measuring trust in AVs using simulator and vehicle platforms that can be extended to other scenario parameters (e.g., varied traffic density, different speed settings, and assertiveness) and events (e.g., pedestrians, cyclists).

References

1. SAE International: U.S. Department of transportation's new policy on automated vehicles adopts SAE International's levels of automation for defining driving automation in on-road motor vehicles (2016). https://www.sae.org/news/3544
2. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**(1), 50–80 (2004)
3. Abe, G., Sato, K., Itoh, M.: Driver's trust in automated driving when passing other traffic objects. In: 2015 IEEE International Conference Systems, Man, and Cybernetics (SMC), pp. 897–902. IEEE (2015)
4. Gold, C., Körber, M., Hohenberger, C., Lechner, D., Bengler, K.: Trust in automation–before and after the experience of take-over scenarios in a highly automated vehicle. Procedia Manufact. **3**, 3025–3032 (2015)
5. Silberg, G., Wallace, R., Matuszak, G., Plessers, J., Brower, C., Subramanian, D.: Self-driving cars: the next revolution. White paper, KPMG LLP & Center of Automotive Research, p. 36 (2012)

6. Morgan, P.L., Alford, C., Williams, C., Parkhurst, G., Pipe, T.: Manual takeover and handover of a simulated fully autonomous vehicle within urban and extra-urban settings. In: 8th International Conference on Applied Human Factors and Ergonomics – Human Factors in Transportation, pp. 760–771. Springer (2017)
7. Bainbridge, L.: Ironies of automation. In: Analysis, Design and Evaluation of Man–Machine Systems, pp. 129–135 (1983)
8. Chambers, A.B., Nagel, D.C.: Pilots of the future: human or computer? Commun. ACM **28** (11), 1187–1199 (1985)
9. Hancock, P.A., Parasuraman, R.: Human factors and safety in the design of intelligent vehicle-highway systems (IVHS). J. Saf. Res. **23**(4), 181–198 (1992)
10. Wiener, E.L., Curry, R.E.: Flight-deck automation: promises and problems. Ergonomics **23** (10), 995–1011 (1980)
11. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse, abuse. Hum. Factors **39**(2), 230–253 (1997)
12. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: Situation awareness, mental workload, and trust in automation: viable, empirically supported cognitive engineering constructs. J. Cogn. Eng. Decis. Making **2**(2), 140–160 (2008)
13. Endsley, M.R.: From here to autonomy: lessons learned from human–automation research. Hum. Factors **59**(1), 5–27 (2017)
14. Hancock, P.A., Billings, D.R., Schaefer, K.E., Chen, J.Y., De Visser, E.J., Parasuraman, R.: A meta-analysis of factors affecting trust in human-robot interaction. Hum. Factors **53**(5), 517–527 (2011)
15. Ghazizadeh, M., Lee, J.D., Boyle, L.N.: Extending the technology acceptance model to assess automation. Cogn. Technol. Work **14**(1), 39–49 (2012)
16. Hoff, K.A., Bashir, M.: Trust in automation: integrating empirical evidence on factors that influence trust. Hum. Factors **57**(3), 407–434 (2015)
17. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **30**(2), 286–297 (2000)
18. Choi, J.K., Ji, Y.G.: Investigating the importance of trust on adopting an autonomous vehicle. Int. J. Hum.-Comput. Interact. **31**(10), 692–702 (2015)
19. Gefen, D., Karahanna, E., Straub, D.W.: Trust and TAM in online shopping: an integrated model. MIS Q. **27**(1), 51–90 (2003)
20. Lee, J.D., Moray, N.: Trust, self-confidence, and operators' adaptation to automation. Int. J. Hum.-Comput. Stud. **40**(1), 153–184 (1994)
21. Bailey, N.R., Scerbo, M.W.: Automation-induced complacency for monitoring highly reliable systems: the role of task complexity, system experience, and operator trust. Theor. Issues Ergon. Sci. **8**(4), 321–348 (2007)
22. Muir, B.M., Moray, N.: Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics **39**(3), 429–460 (1996)
23. Pop, V.L., Shrewsbury, A., Durso, F.T.: Individual differences in the calibration of trust in automation. Hum. Factors **57**(4), 545–556 (2015)
24. Körber, M., Baseler, E., Bengler, K.: Introduction matters: manipulating trust in automation and reliance in automated driving. Appl. Ergon. **66**, 18–31 (2018)
25. Endsley, M.R.: Autonomous driving systems: a preliminary naturalistic study of the tesla model S. J. Cogn. Eng. Decis. Making **11**(3), 225–238 (2017)
26. Ho, G., Wheatley, D., Scialfa, C.T.: Age differences in trust and reliance of a medication management system. Interact. Comput. **17**(6), 690–710 (2005)

27. Sanchez, J., Fisk, A.D., Rogers, W.A.: Reliability and age-related effects on trust and reliance of a decision support aid. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 48, no. 3, pp. 586–589. Sage, Los Angeles (2004)
28. Durso, F.T., Hackworth, C.A., Truitt, T.R., Crutchfield, J., Nikolic, D., Manning, C.A.: Situation awareness as a predictor of performance for en route air traffic controllers. Air Traffic Control Q. **6**(1), 1–20 (1998)
29. Jones, D.G., Endsley, M.R.: Overcoming representational errors in complex environments. Hum. Factors **42**(3), 367–378 (2000)
30. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Factors **37**(1), 32–64 (1995)
31. Endsley, M.R.: Situation awareness misconceptions and misunderstandings. J. Cogn. Eng. Decis. Making **9**(1), 4–32 (2015)
32. Endsley, M.R.: Final reflections: situation awareness models and measures. J. Cogn. Eng. Decis. Making **9**(1), 101–111 (2015)
33. Cohen, J.: A power primer. Psychol. Bull. **112**(1), 155 (1992)
34. Ashalatha, R., Chandra, S.: Critical gap through clearing behavior of drivers at unsignalised intersections. KSCE J. Civ. Eng. **15**(8), 1427–1434 (2011)
35. Mcknight, D.H., Carter, M., Thatcher, J.B., Clay, P.F.: Trust in a specific technology: an investigation of its components and measures. ACM Trans. Manag. Inf. Syst. (TMIS) **2**(2), 12 (2011)
36. Jian, J.Y., Bisantz, A.M., Drury, C.G.: Foundations for an empirically determined scale of trust in automated systems. Int. J. Cogn. Ergon. **4**(1), 53–71 (2000)