Human Interaction and Emerging Technologies (IHIET 2025), Vol. 197, 2025, 399–406
https://doi.org/10.54941/ahfe1006732

Impact of Road Event Recognition Reliability in Autonomous Vehicles on Driver Trust and Takeover Performance

Shan-chih Chen and Yung-ching Liu

Department of Industrial Engineering and Management, National Yunlin University of Science and Technology, Taiwan

ABSTRACT

Autonomous vehicle (AV) recognition reliability is key to building driver trust and ensuring safe handovers. This study examined how different reliability levels (93%, 80%, 60%) and error types (false alarms vs. misses) affect driver trust and takeover behavior using a driving simulator. Sixty participants with no AV experience were divided into two groups: one performed a non-driving task (NDRT), the other passively observed. Over six standardized scenarios, each participant encountered 15 AV recognition events during an 18-minute drive. Results showed that miss errors, especially at lower reliability, significantly reduced trust, increased cognitive load, delayed reaction times in the subsequent takeover task by approximately 0.41 seconds, and led to unstable driving behavior (e.g., sudden lane changes, increased steering variability). In contrast, false alarms had a smaller and non-significant effect, slightly decreasing reaction time in the subsequent takeover task by about 0.14 seconds. After encountering multiple errors, particularly at 80% and 60% reliability, drivers remained skeptical of system alerts. While trust improved slightly with accurate warnings, it never fully recovered to the levels observed with high reliability. The findings highlight the importance of high system reliability and error management in sustaining driver trust and ensuring safe interaction with AVs.

Keywords: Autonomous vehicle (AV), Driver's trust, False alarm, Miss, Reliability, Road event recognition, Takeover

INTRODUCTION

One of the most direct benefits of autonomous driving is its potential to significantly reduce traffic accidents caused by human error by as much as 94% (Singh, 2015). In Taiwan, for example, driver negligence accounted for 96.4% to 98.2% of road traffic accidents (National Police Agency, 2022). This benefit also leads to reduced traffic congestion and emissions, thereby improving overall road efficiency (Choi & Ji, 2015). Most current research on autonomous vehicles(AV) focuses on system accuracy and human-machine interface performance, with less attention given to driver trust in the vehicle (Zhang et al., 2019). However, as fully autonomous systems still lack safety certification, trust remains crucial that is drivers' willingness to relinquish control and respond effectively to emergencies depends heavily on their trust in the system. Therefore, drivers' trust in the vehicle's ability to recognize unexpected road events at varying levels of reliability significantly affects their subsequent response performance.

Trust is one of the core factors influencing a driver's decision to use autonomous driving systems (Bansal, Kockelman & Singh, 2016; Zhang et al., 2019). Manchon (2022) defines trust in autonomous vehicles as "the driver's willingness to delegate driving tasks to the vehicle, even when accident risks exist, in order to enhance safety and comfort." Therefore, drivers must trust the system's reliability to engage with it. This trust is shaped through interaction—whether the system performs well or makes errors—ultimately leading drivers to recalibrate their level of trust (Lee & Moray, 1992; Kraus et al., 2020).

Risk is a key factor influencing trust. In high-risk situations, people are less likely to act on trust, for example, operators may avoid using automation in hazardous environments (Cohen, 2015; Hung et al., 2004). In autonomous driving, uncertainty, such as system errors during sudden road events, can increase perceived risk and reduce trust (Sheehan et al., 2017). When AVs fail to perform reliably, drivers adjust their trust based on perceived risk (Verberne et al., 2012); lower perceived risk generally leads to higher trust (Dinev & Hart, 2006). Thus, understanding how risk perception shapes driver trust is crucial for designing effective AV systems.

System reliability greatly affects trust in automation. Moray et al. (2000) found that when reliability drops below 70%, users perceive the system as untrustworthy, while reliability above 90% stabilizes trust. Studies have used reliability levels ranging from 30% to 100%, with high-risk systems (like AVs) requiring higher reliability thresholds than low-risk ones. In high-risk contexts, reliability below 60% often leads users to abandon the system altogether (de Visser & Parasuraman, 2011).

In highly autonomous vehicles, drivers shift from active controllers to passive observers, changing human-vehicle interaction and making trust a key safety factor. Studies show trust and system monitoring are negatively correlated which is higher trust leads to less monitoring (Muir & Moray, 1996). Drivers with high trust react slower to sudden events, increasing accident risk (Payre et al., 2016). They also spend more time on non-driving-related tasks (NDRTs), reducing road awareness (Petersen et al., 2019), and monitor the system less frequently (Körber et al., 2018). On-road studies found that vehicle speed and NDRT type significantly affect trust and attention. At 30 km/h and 50 km/h, trust was higher when no NDRT was performed compared to visual or manual tasks. Higher speeds improved situational awareness and reduced reaction time (Olaverri-Monreal, 2016).

SAE Level 2–3 AV systems still cannot handle all road situations flawlessly, requiring timely alerts for drivers to take over; NDRTs reduce situational awareness and increase cognitive load, delaying takeover response (Du et al., 2020), but visual-auditory warnings given 10 seconds in advance can help even highly distracted drivers respond effectively (Zeeb, Buchner & Schrauf, 2015).

METHODS

Participants

This study involved 60 participants (30 males, 30 females), aged 20-30, with a valid driver's license and at least weekly driving experience. Participants had normal or corrected vision (0.8 or better) and could engage in conversations. None had prior experience with autonomous driving systems. They were randomly split into two groups: one performed non-driving-related tasks (NDRT), and the other did not. The study was approved by the Human Research Ethics Committee of the National Cheng Kung University Hospital (NCKU IRB No.: A-ER-112-077).

Apparatus

This study utilizes the STI® (STISIM Drive® M1000-R) driving simulator and SDL (Scenario Definition Language 3.22.11) to create experimental scenarios and collect driving behavior data. The experiment is conducted with a VOLVO 340 DL vehicle, equipped with a Logitech G29 DRIVING FORCE steering wheel and pedal system for controlling lateral and longitudinal movements. EDIFIER C3X2.1 speakers deliver experimental sound effects, while an iPad Air 4 tablet is used for participants to perform non-driving related tasks (NDRTs) via E-prime 3.0 cognitive software. Participants are tasked with identifying a single "Q" among 100 letters (99 "O"s) on the tablet and responding by touching it. The tablet is positioned at the vehicle's air conditioning control area, adjustable for the driver's comfort. A 100-point scale questionnaire assesses the driver's trust in the autonomous system.

Driving Scenario

The road scenarios and event scripts required for the experiment were created using SDL (Scenario Definition Language 3.22.11). The experiment includes six road scenarios, based on a 3 (system reliability: 93%, 80%, 60%) × 2 (error type: false alarm vs. miss) design. Each road is 22.2 km in length and takes approximately 18 minutes to complete when driving at the speed limit of 70 km/h. To avoid variations in situational awareness caused by differing environments, all six road scenarios (3 reliability levels × 2 error types) share a consistent driving environment: a 3.65-meter-wide, three-lane, same-direction roadway under daytime and low driving workload conditions.

The experimental scenario begins with manual driving for about 3 minutes, followed by a switch to autonomous mode. No unexpected events occur during manual driving. In autonomous mode, depending on system reliability, a takeover request may or may not be issued, allowing observation of driver response and trust.

The dashboard displays three states: manual driving (Fig. 1(a)), autonomous driving (Fig. 1(b)), and takeover request (Fig. 1(c)). During autonomous mode, participants assigned NDRTs must continue the task when they judge it safe. Takeover alerts use both audio and visual cues: a female voice says "Take over–take over – please switch to manual driving," and a flashing orange-red "Take Over" message appears beside the "Auto" indicator. Drivers must quickly regain control, change lanes to avoid a collision, and the system resumes autonomy once the event is passed.

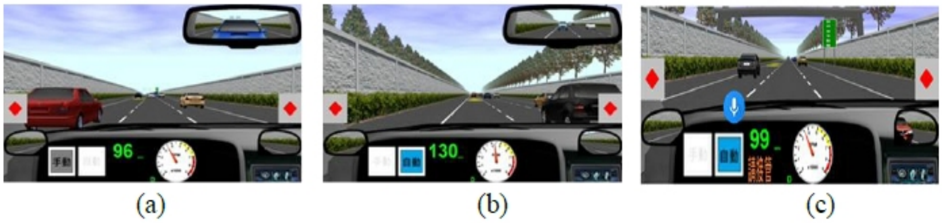

Figure 1: In-vehicle dashboard information display based on vehicle status. (a) When the driver is manually controlling the vehicle, a gray "Manual" icon is shown; (b) when the vehicle is in autonomous mode, a blue "Auto" icon is displayed; (c) when the system encounters an unexpected situation it cannot recognize, the driver is required to take over control. In such cases, the system issues both auditory and visual alerts. The auditory alert is a female voice through the speaker saying, "Please take over - Please take over-switch to manual driving," while the visual alert displays a flashing orange-red "Take over" warning next to the "Auto" icon on the dashboard : voice prompt.

On the autonomous driving route, one event occurred approximately every 1.17 km, or once per minute at the speed limit, totaling 15 events. The event types included: (i) Hit: an unidentified obstacle appears ahead, and the system correctly issues a takeover warning; (ii) Miss: an obstacle is present, but the system fails to detect it and does not issue a warning; (iii) Correct reject: the road ahead is clear, and the system remains silent as expected; (iv) False alarm: there is no obstacle, yet the system incorrectly issues a takeover warning.

The first three events (Events 1–3) and the last three events (Events 13–15) involved only correct system responses (hits or correct rejects) to prevent participants' trust in the autonomous system from being affected by initial or recency effects. Therefore, recognition errors (i.e., false alarms or misses) were arranged to occur between Events 4–12. All events were randomly assigned based on three levels of reliability (93% vs. 80% vs. 60%) and two types of error warnings (miss vs. false alarm). After each event, participants were required to rate their trust in the AV system on a scale from 0 to 100.

Experimental Design and Procedures

This experiment adopted a 3 (Reliability: 93% vs. 80% vs. 60%; within-subject) × 2 (Event error type: False alarm vs. Miss; within-subject) × 2 (Secondary task: Required vs. Not required; between-subject) mixed factorial design. A counterbalanced design was used to control for order effects.

Participants underwent vision and hearing tests, confirmed eligibility, and signed an informed consent form. They then adjusted the controls, practiced with the simulator, and familiarized themselves with the tasks and error types.

After practice, they completed a trust questionnaire before starting the formal experiment.

The participants were divided into two groups (with or without NDRT), and the experiment consisted of six road segments, each lasting 18 minutes. To avoid fatigue, the experiment was split into two sessions, with partial compensation given after the first session. Sessions were scheduled based on participants' availability, with at least one day between them. Remaining compensation was given after the second session.

The experimental data collection includes drivers' takeover reaction time, driving behavior, and level of trust in the autonomous driving system.

RESULTS

There was a significant interaction between system reliability and secondary task execution on drivers' average trust ($F(2, 122) = 10.441$, $p < .001$). As shown in Figure 2(a), at both 80% and 60% reliability levels, performing a secondary task significantly decreased trust compared to not performing one ($p = .016$ and $p < .001$, respectively).

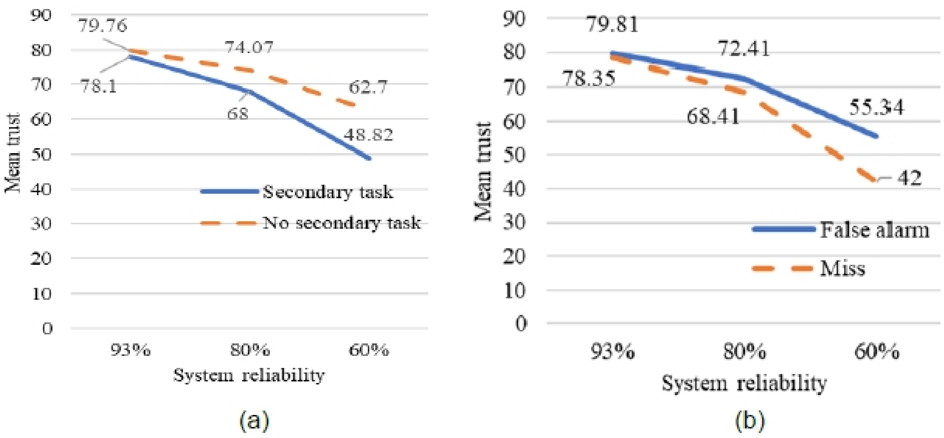

Figure 2: Interaction effect on drivers' average trust between (a) system reliability and secondary task execution; and (b) between system reliability and event error type.

Additionally, a significant interaction was found between system reliability and event error type ($F(2, 122) = 11.641$, $p < .001$). At both reliability levels, event error type significantly affected trust ($F(1, 59) = 11.707$, $p = .001$; $F(1, 59) = 27.474$, $p < .001$), with false alarms eliciting higher trust than misses (Figure 2(b)).

Results showed that encountering a miss error significantly delayed drivers' reaction time in a subsequent takeover task by approximately 0.41 seconds ($t(57) = 2.806$; $p = 0.007$). In contrast, experiencing a false alarm led to a slight, non-significant decreased in takeover reaction time during the next event by about 0.14 seconds ($t(58) = -0.932$; $p = 0.355$). In takeover time performance, hit and miss events represent two extreme conditions. Post-hoc analysis revealed that, overall, drivers' takeover times were significantly slower following miss events compared to hit events by an average of 0.65 seconds ($p < .001$). Notably, when system reliability was at 93% and drivers were engaged in an in-vehicle secondary task, this difference increased to a striking 1.57 seconds (hit vs miss: 4.5212 sec vs. 6.0914 sec).

Driving behavior was influenced by road events, with miss events resulting in the greatest lateral acceleration variability, followed by hit and false alarm (FA) events, respectively ($F(2, 110) = 64.616$, $p <.001$). The results indicate that miss errors elicited unprepared and destabilized responses from drivers.

DISCUSSION

This study examined how system reliability, secondary task engagement, and error type interact to affect drivers' trust and behavior in automated driving. Results show that trust is highly context-dependent, decreasing significantly when drivers perform secondary tasks under lower reliability conditions, consistent with findings that cognitive load amplifies sensitivity to automation errors (Dzindolet et al., 2003).

Error type also played a key role: false alarms consistently elicited higher trust than misses, supporting prior research showing that omission errors (misses) are perceived as more dangerous and more damaging to user confidence than commission errors (Lee & See, 2004). Behaviorally, misses led to significantly slower takeover reaction times and greater lateral acceleration variability, indicating impaired readiness and control—especially under multitasking conditions. This aligns with earlier work showing that critical event detection delays and instability tend to follow trust-disrupting events like system misses (Zeeb et al., 2015).

Overall, miss errors not only eroded trust but also compromised safety by delaying driver response and destabilizing control. These findings suggest that trust in automation hinges not just on reliability, but on real-time interactions—highlighting the need for transparent system feedback and robust error-handling mechanisms in distracting or high-load environments.

ACKNOWLEDGMENT

The authors express gratitude to the National Science and Technology Council (NSTC) for funding this research project under grant number 112-2221-E-224-034.

REFERENCES

Bansal, P., Kockelman, K. M., & Singh, A. (2016). Assessing public opinions of and interest in new vehicle technologies: An Austin perspective. Transportation Research Part C: Emerging Technologies, 67, 1–14.

Choi, J. K., Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. International Journal of Human-Computer Interaction, 31(10), 692–702.

de Visser, E., & Parasuraman, R. (2011). Adaptive aiding of human-robot teaming: Effects of imperfect automation on performance, trust, and workload. Journal of Cognitive Engineering and Decision Making, 5(2), 209–231.

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. Information systems research, 17(1), 61–80.

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. International Journal of Human-Computer Studies, 58(6), 697–718. https://doi.org/10.1016/S1071-5819(03)00038-7.

Du, N., Kim, J., Zhou, F., Pulver, E., Tilbury, D. M., Robert, L. P., Yang, X. J. (2020). Evaluating effects of cognitive load, takeover request lead time, and traffic density on drivers' takeover performance in conditionally automated driving. Paper presented at the 12th international conference on automotive user interfaces and interactive vehicular applications.

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

Körber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. Applied ergonomics, 66, 18–31.

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. Ergonomics, 35(10), 1243–1270.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human Factors, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392.

Manchon, J.-b., Bueno, M., & Navarro, J. (2022). How the initial level of trust in automated driving impacts drivers' behaviour and early trust construction. Transportation research part F: Traffic psychology and behaviour, 86, 281–295.

Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. Journal of experimental psychology: Applied, 6(1), 44.

Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics, 39(3), 429–460.

National Police Agent (June, 2022). Website: https://www.npa.gov.tw/ch/app/data/view?module=wg056&id=2217&serno=fbf5a752-9c50-4eb1-a3ac-8d4f7112c70a

Olaverri-Monreal, C. (2016). Autonomous vehicles and smart mobility related technologies. Infocommunications Journal, 8(2), 17–24.

Payre, W., Cestac, J., & Delhomme, P. (2016). Fully automated driving: Impact of trust and practice on manual control recovery. Human factors, 58(2), 229–241.

Petersen, L., Robert, L., Yang, X. J., & Tilbury, D. M. (2019). Situational awareness, drivers trust in automated driving systems and secondary task performance. arXiv preprint arXiv:1903.05251.

Sheehan, B., Murphy, F., Ryan, C., Mullins, M., & Liu, H. Y. (2017). Semi-autonomous vehicle motor insurance: A Bayesian Network risk transfer approach. Transportation Research Part C: Emerging Technologies, 82, 124–137.

Singh, S. (2015). Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey. Traffic Safety Facts Crash•Stats. Report No. DOT HS 812 115. Washington, DC: National Highway Traffic Safety Administration.

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. Accident analysis & prevention, 78, 212–221.

Zhang, T., Tao, D., Qu, X., Zhang, X., Lin, R., & Zhang, W. (2019). The roles of initial trust and perceived risk in public's acceptance of automated vehicles. Transportation research part C: emerging technologies, 98, 207–220.