Proceedings of the Human Factors and Ergonomics Society Annual Meeting
2024, Vol. 68(1) 25–26

DOI: 10.1177/10711813241280937
journals.sagepub.com/home/pro
Sage

Eye, Heart, the Brain: The Psychophysiology of Trust in AVs

Yinsu Zhang[1], Ranjana K. Mehta[1], and Anthony D. McDonald[1]

Abstract

Automation misuse and acceptance, influenced by trust, environmental conditions, and confidence, have hindered drivers from fully benefiting from partially automated vehicles. This study investigates how driver trust changes with AV reliance, differences in mental and physiological states, and continuous measures' effectiveness. The takeover drivers reported lower trust than the non-takeover drivers in all scenarios. Nontakeover drivers' elevated DLPFC activation aligns with trust networks and emotion regulation. The groups also differed in neural activation preand during scenarios with the takeover group showed more PFC, V2V3, and IFC engagement pre-scenario. Gaze revealed the takeover group fixated more on the AV button or dashboard, indicating readiness to take over, while non-takeover drivers focused on the rearview mirror, reflecting situational awareness. HRV responses showed higher physiological arousal in the takeover group pre-scenario. In summary, our multimodal approach reveals takeover behavior is associated with lower trust, cognitive unloading, increased stress, and anticipatory visual attention.

Keywords

trust, automated vehicles, fNIRS, eye tracking, heart rate variability, neural correlates, fatigue, workload, situation awareness

Since the introduction of partially automated vehicles (AVs), automation misuse and technology acceptance have been preventing drivers from realizing AV's full advantages. Various factors, such as trust, environmental conditions, and self-confidence, determine the extent to which drivers rely on automated systems (Gao & Lee, 2006).

Trust in automation significantly influences driver interactions with AVs (Lee & See, 2004). Continuous evaluation of trust in AVs is necessary, as trust is a dynamic and latent concept. Popular trust surveys are discrete and may not fully capture drivers' dynamic cognitive processes (Jahedi & Méndez, 2014). Objective measures of neural and physiological responses offer the potential to complement trust surveys and capture dynamic changes in trust.

Studies have examined neural correlates of driver states: Increased activity in the prefrontal cortex has been observed during manual driving compared to partially automated driving (Sibi et al., 2016). Also, distrust in AVs is associated with heightened activation in the Broca's area, inferior frontal cortex (IFC), and the dorsal-lateral prefrontal cortex (DLPFC) (Seet et al., 2022). Eye-tracking fixations have also been proven effective in capturing drivers' perceptions and behavioral responses to automation (Rudin-Brown & Noy, 2004). Gaze behaviors of stationary gaze entropy (SGE) and gaze transition entropy (GTE) have been linked to fundamental driver states and cognition (Chen et al., 2022; Shiferaw et al., 2018). Physiological signals, such as heart rate variability (HRV), have been established as reliable indicators of mental workload, stress, and fatigue (Nacpil et al., 2021). Changes in HRV have been associated with variations in trust levels, suggesting its potential to predict driver takeover behaviors (Du et al., 2020).

This study aims to explore the perception-cognition-action processes to answer how driver trust change based on AV reliance decisions, the differences in mental and physiological states among drivers with varying takeover behaviors, and the effectiveness of continuous measures in reflecting trust. Fifty-seven drivers participated in a single session of AV interactions in a medium fidelity driving simulator. All had valid US driver's licenses and no prior AV experience. They were allowed to initiate takeovers at any point during the experiment, reflecting their actual reliance behavior, and were instructed to engage the AV system when safe. The experiment included four variable AV performance scenarios: two crash avoidance and two silent failure scenarios. In-drive surveys included the Situational Trust Scale for Automated Driving (STS-AD) (Holthausen et al., 2020), the Situational Awareness Rating Technique (SART) survey (Taylor, 1989),

[1]University of Wisconsin Madison, WI, USA

Corresponding Author:
Yinsu Zhang, University of Wisconsin Madison, Mechanical Engineering Building, 3107, 1513 University Ave, Madison, WI 53706-1314, USA.
Email: yinsu.zhang@wisc.edu

and single-item questions on fatigue and workload. Mean and peak neural activations were collected using functional near-infrared spectroscopy (fNIRS). HRV data were analyzed for indicators of fatigue and stress. Gaze behavior was tracked for fixations, SGE, and GTE to uncover cognitive processes.

Linear mixed-effect models were utilized with participants as random factors to assess differences between phases (pre- and during scenarios), and behaviors (based on participant non-takeover and takeover behaviors). Independent samples *t*-tests evaluated differences in subjective responses and takeover behaviors.

Across scenarios, the takeover group perceived lower trust in AVs, and the non-takeover drivers' higher overall trust levels were primarily driven by aspects of AV performance, judgment, and risk perceptions. Heightened neural activations of non-takeover drivers in DLPFC corroborate with existing evidence on trust networks, modulated with emotion regulation. Causal studies have shown that excitatory DLPFC activity is related to enhanced trust (Hopko & Mehta, 2022). We also found that the two groups differed in their neural activation pre- versus during scenarios. Consistently we observed that the takeover group exhibited greater engagement of the PFC, V2V3, and IFC pre-scenario than during the scenario, while the reverse was found for the non-takeover group. Such brain dynamics align with findings from studies where lower trust correlated with higher prefrontal cortex activation (Hopko & Mehta, 2022; Seet et al., 2022), suggesting the top-down process of anticipatory attentional control.

The takeover group fixated more on the AV button or dashboard areas and exhibited exploratory gaze behaviors, indicating their readiness to take over manual control and increased uncertainty. In contrast, the non-takeover drivers tended to focus more on specific areas like the rearview mirror, indicating reflective behavior and heightened situational awareness necessary for the top-down process involved. The increased anticipatory attentional control in the takeover group pre-scenario was also evident in HRV responses, wherein the group exhibited greater physiological arousal pre-scenario when compared to during the scenario.

Our multimodal approach depicts associations of takeover behavior with lower subjective trust and cognitive unloading, and higher physiological stress and anticipatory visual attention.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partially supported by the U.S. National Science Foundation (Award#: 2310621).

ORCID iDs

Yinsu Zhang https://orcid.org/0000-0002-0954-0973
Anthony D McDonald https://orcid.org/0000-0001-7827-8828

References

Chen, W., Sawaragi, T., & Hiraoka, T. (2022). Comparing eye-tracking metrics of mental workload caused by NDRTs in semi-autonomous driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *89*, 109–128. https://doi.org/10.1016/j.trf.2022.05.004

Du, N., Yang, X. J., & Zhou, F. (2020). Psychophysiological responses to takeover requests in conditionally automated driving. *Accident Analysis & Prevention*, *148*, 105804. https://doi.org/10.1016/j.aap.2020.105804

Gao, J., & Lee, J. D. (2006). Extending the decision field theory to model operators' reliance on automation in supervisory control situations. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *36*(5), 943–959. https://doi.org/10.1109/TSMCA.2005.855783

Holthausen, B. E., Wintersberger, P., Walker, B. N., & Riener, A. (2020). Situational trust scale for automated driving (STS-AD): Development and initial validation. *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 40–47). https://doi.org/10.1145/3409120.3410637

Hopko, S. K., & Mehta, R. K. (2022). Trust in shared-space collaborative robots: Shedding light on the human brain. *Human Factors, 66*(2), 490–509. https://doi.org/10.1177/00187208221109039

Jahedi, S., & Méndez, F. (2014). On the advantages and disadvantages of subjective measures. *Journal of Economic Behavior & Organization*, *98*, 97–114. https://doi.org/10.1016/j.jebo.2013.12.016

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

Nacpil, E. J. C., Wang, Z., & Nakano, K. (2021). Application of physiological sensors for personalization in semi-autonomous driving: A review. *IEEE Sensors Journal*, *21*(18), 19662–19674. https://doi.org/10.1109/JSEN.2021.3100038

Rudin-Brown, C., & Noy, I. Y. (2004). Behavioural adaptation to in-vehicle safety measures: Past ideas and future directions. In T. Rothengatter & E. C. Vava (Eds.), *Traffic and transport psychology: Theory and application* (pp. 25–46). Elsevier.

Seet, M., Harvy, J., Bose, R., Dragomir, A., Bezerianos, A., & Thakor, N. (2022). Differential impact of autonomous vehicle malfunctions on human trust. *IEEE Transactions on Intelligent Transportation Systems*, *23*(1), 548–557. https://doi.org/10.1109/TITS.2020.3013278

Shiferaw, B. A., Downey, L. A., Westlake, J., Stevens, B., Rajaratnam, S. M. W., Berlowitz, D. J., Swann, P., & Howard, M. E. (2018). Stationary gaze entropy predicts lane departure events in sleep-deprived drivers. *Scientific Reports*, *8*(1), 2220. https://doi.org/10.1038/s41598-018-20588-7

Sibi, S., Ayaz, H., Kuhns, D., Sirkin, D., & Ju, W. (2016). Monitoring driver cognitive load using functional near infrared spectroscopy in partially autonomous cars. In *2016 IEEE intelligent vehicles symposium (IV)* (pp. 419–425). https://doi.org/10.1109/IVS.2016.7535420

Taylor, R. M. (1989). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. Proceedings of the AGARD AMP symposium on situational awareness in aerospace operations, CP478.