Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira,
Edoardo Patelli, Hélder S. Sousa, Rui A. Silva, Monica Santamaria and Chiara Turco
© 2026 ESREL 2026 Organizers. *Published by* Research Publishing, Singapore.
ISBN: 978-981-94-6718-1 | doi: 10.3850/ESREL2026061419_esrel26-p26367-cd

Effects of Trust in Automation and Takeover System Type on Driver Takeover Behavior in Level 3 Conditionally Automated Driving: A human-in-the-loop driving simulation study

Yaonan Ding

School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China. E-mail: dingyaonan@buaa.edu.cn

Shengkui Zeng

School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China. E-mail: zengshengkui@buaa.edu.cn

Jianbin Guo*

School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China. E-mail: guojianbin@buaa.edu.cn

Haiyang Che

School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China. E-mail: chehaiyang@buaa.edu.cn

Abstract: Level 3 conditionally automated driving permits engagement in Non-Driving Related Tasks, positioning driver trust as a critical factor that modulates monitoring behavior and takeover readiness. Although graded warning strategies (e.g., two-stage warnings incorporating a Monitoring Request) show promise in facilitating takeover transitions, existing research has largely examined warning design and trust in isolation. Consequently, the interactive mechanism between advanced graded warning strategies and driver trust states remains underexplored. This study conducted a driving simulation experiment utilizing a 2 (Trust: High vs. Low) × 2 (Warning: Single-stage vs. Two-stage) within-subjects design to evaluate the effects of these variables on takeover reaction time and collision rates. Results revealed a distinct contrast between response speed and safety outcomes. While both the two-stage warning condition and low trust levels significantly accelerated takeover reaction times, the safety benefits of the warning system were strictly modulated by trust levels. Specifically, statistical analysis indicated a significant interaction: low-trust drivers achieved the highest avoidance success rates under the two-stage condition, whereas high-trust drivers exhibited a significant decline in safety performance under the same condition, despite faster reactions. These findings highlight an automation paradox: excessive trust can be counterproductive specifically when coupled with advanced graded warnings. The study suggests that the performance degradation in high-trust drivers may be attributed to a confirmation bias, where the pre-warning signal is misinterpreted as a routine status update. Consequently, the study advocates for adaptive takeover systems that dynamically calibrate warning strategies based on the driver's trust state to ensure the efficacy of human-machine interaction.

***Keywords*:** conditionally automated driving, trust in automation, takeover warning systems, takeover performance, driving safety.

1. Introduction

The commercialization of Level 3 (L3) conditionally automated driving marks a significant paradigm shift in surface transportation (Gold et al. 2018). By permitting drivers to physically and cognitively disengage from the Dynamic Driving Task (DDT) to engage in Non-Driving Related Tasks (NDRTs), L3 automation holds the potential to significantly enhance traffic efficiency and mobility comfort (Tan and Zhang 2025a). However, this technological advancement fundamentally transforms the human role from an active, continuous operator to a passive, fallback-ready user (Chen et al. 2024; Shahini and Zahabi 2022). This transition exemplifies the Irony of Automation: as system reliability increases, the driver's situation awareness (SA) is

prone to degradation due to the Out-of-the-Loop (OOTL) phenomenon (Hu et al. 2024). Consequently, when the system reaches its Operational Design Domain (ODD) limits, the re-integration of a disengaged driver via a Take-Over Request (TOR) constitutes a critical point of vulnerability, where failure to rapidly reconstruct SA can result in catastrophic collisions.

To mitigate the performance degradation associated with OOTL, the takeover warning system (TWS) serves as the principal intervention. Despite the emergence of graded warning strategies in advanced systems, the conventional single-stage strategy remains the dominant paradigm in empirical research. However, this abrupt-onset approach faces an inherent timing dilemma: extensive lead times may exceed the detection capabilities of sensors (e.g., radar accuracy), potentially inducing false alarms that undermine driver trust. Conversely, short lead times often provoke a startle effect, leaving sufficient time for both SA restoration and motor preparation. To address these limitations, graded warning strategies—specifically implemented as two-stage warning systems incorporating a preliminary, lower-urgency Monitoring Request (MR)—have been proposed as a superior alternative. Theoretically, the MR functions as a cognitive scaffold, guiding driver attention progressively from the NDRT back to the driving environment (Agrawal and Peeta 2021). Crucially, it also facilitates motor readiness, prompting drivers to reposition their hands and feet before the critical limit is reached, thereby ensuring a smoother transition of control.

Despite the theoretical advantages of graded warnings, the effectiveness of any TWS is not determined solely by its physical design; rather, it is strictly modulated by the driver's internal cognitive state, particularly trust in automation (Tan and Zhang 2025b). Trust is widely regarded as a pivotal determinant of human-machine collaboration and serves as the fundamental basis for driver-automation interaction (Qu et al. 2023).

Figure 1 illustrates the trust calibration framework. Deviations from the optimal diagonal undermine system utility: under-trust (disuse) negates L3 efficiency benefits, while over-trust (misuse) breeds complacency. Crucially, the latter compromises the effectiveness of advanced warnings, exposing a latent safety hazard in current standardized alerting strategies.

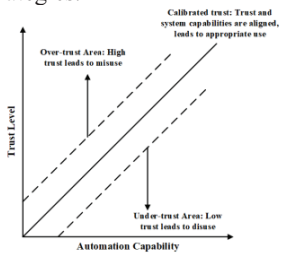


Figure 1. Relationship among trust levels and automation capability (Papadimitriou et al. 2020).

While the critical roles of warning design and trust in automation are individually recognized, existing literature has largely examined these factors in isolation. Existing literature has extensively optimized warning attributes—ranging from timing and modality to semantic content—to enhance takeover performance (i.e., timeliness and quality). Parallelly, research on trust calibration has well documented how automation complacency impairs monitoring behaviors and subsequent intervention quality. However, despite the emerging adoption of multi-stage warning strategies, the interactive mechanism between these advanced strategies and trust in automation remains underexplored. Specifically, prior studies have predominantly evaluated graded warnings under nominal or controlled trust conditions. Consequently, it remains an open question whether the safety benefits of such systems are robust across varying levels of driver trust, or if their efficacy is subject to boundary conditions imposed by the driver's cognitive state.

To bridge this gap, this study employs a high-fidelity driving simulation to systematically investigate the interplay between warning condition and trust level on takeover performance. Moving beyond isolated assessments, our objective is to determine whether the safety efficacy of TWSs is contingent upon the driver's trust level, thereby informing the design of adaptive TWSs. We propose the following hypotheses:

H1: Drivers receiving the two-stage warning will exhibit significantly shorter takeover reaction times compared to those receiving the single-stage warning.

H2: Drivers with low trust in automation will demonstrate faster takeover reaction times compared to drivers with high trust levels.

H3: We hypothesize that the safety benefits of two-stage warnings are significantly moderated by driver trust, yielding a significant interaction effect.

2. Related Work

2.1. Takeover Warning Systems and Takeover Performance

Existing research has extensively optimized the design parameters of conventional single-stage TORs across three critical dimensions (Deng et al. 2024): **(1) Temporal:** Establishing time budgets (typically 3–7s) to accommodate perception-reaction times; **(2) Modal:** Utilizing multimodal alerts (visual-auditory-haptic) to accelerate sensory processing; and **(3) Semantic:** Conveying explanatory information (e.g., takeover reasons) to minimize decision errors.

Despite these parameter optimizations, single-stage strategies remain limited by their abrupt onset. The sudden demand for control not only induces a startle but also frequently fails to provide a sufficient time budget for the sequential processes of cognitive reconstruction and physical motor preparation (Chen et al. 2025). To address these deficits, two-stage warning strategies incorporating a preliminary MR have been proposed. Functioning as a cognitive scaffold or uncertainty notification, the MR guides attention back to the driving environment prior to the critical limit (Lu et al. 2019), thereby reducing mental workload and enhancing takeover readiness. However, the efficacy of such staged alerts relies on driver compliance; frequent issuance of low-urgency MRs may trigger a "cry wolf" effect, potentially desensitizing drivers to system prompts (Xu et al. 2022).

2.2. Trust Calibration and Behavioral Adaptation

Trust in automation is recognized as the fundamental basis of effective Human-Automation collaboration, acting as the primary mediator determining driver reliance strategies (Yin et al. 2021). Driver behavior is governed by trust calibration—the alignment of subjective trust with objective reliability. Under-trust (where subjective trust falls below objective reliability) typically induces system rejection and unnecessary mental workload (Hu et al. 2023).

Conversely, over-trust (where subjective trust exceeds objective reliability) suppresses vigilance, shifting attention toward NDRTs and leaving drivers unprepared for takeovers (Petersen et al. 2019). Such complacency delays hazard recognition and prolongs reaction times to system failures (Borowsky et al. 2022), while also causing drivers to overlook system cues, heightening the risk of misuse.

2.3. Summary and Research Gap

Reviewing the literature on warning design and trust calibration reveals a critical theoretical misalignment. On one hand, research on TWS advocates for two-stage strategies, assuming that the preliminary MR effectively acts as a cognitive scaffold to guide attention. On the other hand, trust research indicates that high trust naturally elevates the response threshold to system alerts, making drivers less likely to attend to low-urgency signals.

Consequently, the interplay between these two domains remains underexplored. Given the limited empirical evidence regarding the robustness of two-stage warnings against high-trust-induced vigilance decrements, it is unknown whether high trust undermines the MR's intended function, thereby offsetting the advantages of the graded warning design.

3. Methods

3.1. Apparatus and Participants

The study utilized a high-fidelity fixed-base simulator based on Carla (Town04), integrated with a Logitech G923 system and triple screens (Figure 2). 14 licensed drivers with normal vision were recruited. Written informed consent was obtained from all participants.



Figure 2. The experimental setup.

3.2. Experimental Design and Tasks

A 2 (Warning: Single-stage vs. Two-stage) × 2 (Trust: High vs. Low) within-subjects factorial

design was employed to investigate the interactive effects of warning conditions and trust levels on takeover performance. Participants completed four experimental drives, with the order counterbalanced. To simulate the real-world trade-off between safety and efficiency, a performance-based compensation scheme was implemented: it heavily penalized collisions (safety priority), rewarded NDRT progress, and marginally weighted vigilance performance.

The ego-vehicle operated in an advanced Adaptive Cruise Control (ACC) mode with a set speed of 60 km/h. Longitudinal speed was dynamically adjusted based on both lead traffic and road geometry (e.g., curvature) to ensure safety, while lane changing was strictly prohibited. The scenario included: **(1) NDRT:** Playing a puzzle game ("2048") on a handheld smartphone to induce an OOTL state; **(2) Vigilance Task:** Responding to visual probes, designed to mimic naturalistic intermittent monitoring and maintain baseline arousal; and **(3) Takeover Task:** Comprising both System Failures (requiring immediate intervention, e.g., stationary obstacles) and False Alarms (requiring no action).

3.3. Independent Variables

Warning Condition: Single-stage: Multimodal TOR at TTC=5s. Two-stage: Auditory-only MR at TTC=10s (cognitive priming), followed by TOR at TTC=5s.

Trust Level: Manipulated via a hybrid induction paradigm. Prior to each drive, high trust involved an 80% reliability instruction and 5 successful avoidances; low trust involved a 20% reliability instruction and 5 failures. Subjective trust was assessed using the 12-item Automation Trust Scale (Jian et al. 2000). Participants rated their agreement on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).

3.4. Procedure and Data Analysis

Participants completed four experimental drives, with the order counterbalanced across participants to control for sequence effects. A 5-minute mandatory rest break was enforced between drives to minimize carryover effects. To introduce uncertainty, each drive contained 7 events (4 System Failures + 3 False Alarms) presented in a randomized order, separated by intervals of approximately 2 minutes.

Performance metrics were derived exclusively from System Failure events. Dependent variables included Takeover Reaction Time (RT) (interval from TTC=5s to the first effective input: steering >2° or braking >10%) and Collision Occurrence (binary: Collision vs. Non-collision). Analyses were performed using IBM SPSS Statistics 27.0 (alpha = .05).

4. Results

4.1. Manipulation Checks

To assess the efficacy of the trust manipulation, a 2 (Warning: Single vs. Two) × 2 (Trust: High vs. Low) repeated measures ANOVA was conducted on the subjective trust ratings. The analysis yielded a statistically significant main effect of Trust ($F(1, 13) = 96.49$, $p < .001$, $\eta^2 = .881$). As intended by the experimental design, participants reported significantly higher trust in the high-trust condition ($M = 4.24$, $SE = 0.10$) compared to the low-trust condition ($M = 3.25$, $SE = 0.07$).

A significant main effect of warning condition was also observed ($F(1, 13) = 31.54$, $p < .001$, $\eta^2 = .708$), with the two-stage condition eliciting higher trust ratings than the single-stage condition. However, and most importantly for the manipulation check, the interaction between Trust and Warning was not significant ($F(1, 13) = 1.79$, $p = .204$, $\eta^2 = .121$). This nonsignificant interaction confirms that the manipulation of trust levels was robust and remained effective across different warning strategies.

4.2. Takeover Reaction Time

A Linear Mixed Model (LMM) was conducted to analyze takeover reaction times, with warning condition and trust level as fixed effects, subject ID as a random effect, and takeover event order as a covariate. The results are summarized in Table 1. There was a significant main effect of warning condition ($F(1, 111.57) = 10.91$, $p = .001$). Specifically, the mean reaction time in the two-stage warning condition ($M = 2.79$s, $SE = 0.12$) was significantly shorter than that in the single-stage condition ($M = 3.17$s, $SE = 0.12$), representing a reduction of 0.38 seconds (Figure 3).

Table 1. Results of Linear Mixed Model (LMM) Analysis for RT.

Effect	*df*	*F*	*p*
EventID	1, 113.73	.09	.763
Warning	1, 111.57	10.91	.001**
Trust	1, 111.94	19.15	**<.001*** **
Warning × Trust	1, 112.16	.047	.493

Note. *df* = degrees of freedom (numerator, denominator). * $p < .05$. ** $p < .01$. *** $p < .001$.

Additionally, a significant main effect of trust level was observed ($F(1, 111.94) = 19.15$, $p < .001$). Drivers with low trust levels exhibited significantly faster reaction times ($M = 2.74$ s, $SE = 0.12$) compared to those with high trust levels ($M = 3.22$ s, $SE = 0.12$), with a mean difference of 0.48 seconds.

However, the interaction effect between warning condition and trust level was not significant ($F(1, 112.16) = 0.47$, $p = .493$), indicating that the benefit of the two-stage warning on reaction speed was consistent across trust levels. The covariate, event order, was not significant ($p = .763$), suggesting no confounding practice effects.

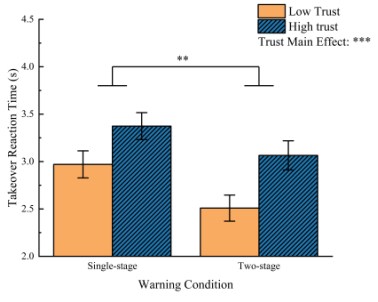


Figure 3. Mean RTs across warning conditions and trust levels. Error bars represent standard errors. (** $p < .01$, *** $p < .001$).

4.3. Takeover Success Rate

To evaluate safety performance, collision data were analyzed by modeling the probability of takeover success (1 = non-collision; 0 = collision) using a Generalized Linear Mixed Model (GLMM) with a binomial distribution and a logit link function. The model included Trust and Warning as fixed effects, and Subject ID as a random effect to account for within-subject correlations. As shown in Table 2, the Type III tests of fixed effects indicated that neither the main effect of warning condition ($F(1, 165) = 0.06$, $p = .815$) nor the main effect of trust level ($F(1, 165) = 2.81$, $p = .096$) was statistically significant.

Table 2. Results of Generalized Linear Mixed Model (GLMM) Analysis for Takeover Success Rate.

Effect	*df*	*F*	*P*
Warning	1, 165	0.06	.815
Trust	1, 165	2.81	.096
Warning × Trust	1, 165	4.21	**.042***

Note. *df* = degrees of freedom (numerator, denominator). * $p < .05$.

However, a significant interaction effect between warning condition and trust level was observed ($F(1, 165) = 4.21$, $p = .042$). To decompose this interaction, simple effects analyses were performed (as shown in Figure 4). The results revealed distinct patterns:

Under the single-stage warning condition, there was no significant difference in avoidance success rates between trust levels ($p > .05$), with both groups maintaining relatively high performance.

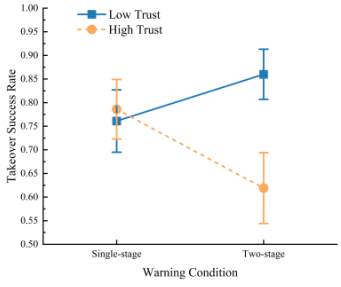


Figure 4. Interaction effect of warning condition and trust level on takeover success rate. Error bars represent standard errors (SE).

In contrast, under the two-stage warning condition, trust level significantly predicted safety outcomes ($p = .011$). Specifically, drivers with low trust demonstrated a significantly higher success rate compared to those with high trust. Parameter estimates indicated that under the two-stage system, the odds of successful avoidance for low-trust drivers were 4.36 times higher than for high-trust drivers (OR = 4.36, 95% CI [1.42, 13.42]).

5. Discussion

5.1. The Temporal Benefit of Two-Stage Warnings: Attention Management and Preparation

The results of this study confirm that the two-stage warning strategy significantly reduces drivers' takeover reaction times. This finding fully supports Hypothesis 1, validating the

temporal advantage of graded warning strategies. Analyzed within the framework of attentional resource allocation, single-stage warnings necessitate an abrupt reorientation of attention from NDRTs to the driving task. This abrupt interruption imposes a substantial switch cost and high instantaneous cognitive load.

In contrast, the two-stage warning strategy mitigates this conflict by introducing the MR as a pre-warning signal. The MR intervention interrupts the driver's focus on the NDRT, effectively mitigating attentional lock-in and offering a temporal buffer prior to the critical TOR. This graded guidance mechanism promotes a progressive transition from an out-of-the-loop to an in-the-loop state. Consequently, upon the activation of the final TOR, the driver is already primed with an elevated level of SA arousal, enabling an expedited motor response.

5.2. The Cost of High Trust: Vigilance Decrement and Response Latency

The statistical results regarding RT fully support Hypothesis 2, confirming the potential negative impact of high trust levels on takeover performance. While trust serves as a foundation for human-machine collaboration, the data indicate that excessive trust levels significantly prolong drivers' RTs. This finding corroborates the established phenomenon of automation complacency.

This prolongation of reaction time can be attributed to two cognitive processes induced by high trust. The first is monitoring inefficiency. Consistent with complacency theory, high-trust drivers tend to perceive the system as highly reliable, leading to a diversion of cognitive resources from monitoring the driving environment in favor of engaging in NDRTs. This shift in attention results in a perceptual delay in detecting warning signals. The second is decision latency induced by over-reliance. Even upon perceiving the warning, high-trust drivers may remain subject to an expectancy bias that the system is capable of handling the situation. This erroneous expectation leads to hesitation in the intervention decision, translating into an active postponement of the takeover. These results suggest that over-trust undermines the driver's vigilance, posing a potential safety risk in critical scenarios.

5.3. The Conditional Efficacy of Two-Stage Warnings: Signal Interpretation and Trust Calibration

Unlike reaction time, safety outcomes showed a complex interaction pattern rather than significant main effects. This absence of main effects is partially attributable to a ceiling effect induced by the predictable experimental scenario, which masked overall group differences. However, the significant interaction reveals that safety risks emerge specifically when the warning strategy is mismatched with the driver's trust level.

This interaction highlights the modulating role of trust in signal interpretation. For low-trust drivers, the MR functioned as a cognitive scaffold that validated their appropriate distrust. This mechanism facilitated SA restoration, prompting them to fully utilize the pre-warning window for preparation. Conversely, for high-trust drivers driven by confirmation bias, the MR was likely misinterpreted as a routine status update rather than a precursor to an impending failure. This erroneous expectation resulted in a failure to utilize the preparation opportunity, leading to a significant performance decline upon the subsequent TOR.

These findings expose an Automation Paradox: while the vision of automated driving is to cultivate trust to relieve drivers from the DDT, our data suggest that in critical takeover phases, high trust can be counterproductive when coupled with two-stage warnings. This implies that future TWS designs must dynamically calibrate driver trust to ensure the safety benefits of graded warning strategies are realized.

5.4. Implications

The interaction effects observed in this study establish two critical design principles for L3 automated driving.

First, implementation of adaptive warning protocols. High trust undermined the efficacy of the low-urgency MR, as complacent drivers failed to initiate preparation. This necessitates a dynamic warning strategy: upon detecting excessive trust, the system should amplify the sensory salience of the pre-warning or adopt directive interaction modalities (e.g., explicit voice commands). These adjustments are essential to counteract confirmation bias and enforce rapid attentional re-engagement.

Second, enhancing signal distinctiveness. The failure of high-trust drivers to respond to the MR suggests that current graded warnings lack sufficient distinctiveness to compete with their expectancy bias. To prevent the MR from being misinterpreted as a routine status update, TWS design must distinctively decouple the coding logic of system status notifications from takeover requests. By explicitly visualizing system confidence levels (transparency), the design can force a calibration of the driver's mental model, ensuring the MR is correctly interpreted as a trigger for SA restoration.

5.5. Limitations and Future Research

This study acknowledges three specific limitations warranting future investigation.

First, the limited scenario complexity—characterized by static obstacles and predictable time budgets—resulted in uniformly high avoidance success rates. This lack of task variability may have masked specific warning effects. Future research should introduce dynamic and unpredictable hazards to rigorously test the efficacy of two-stage warnings under higher workload conditions.

Second, distinct trust states were induced via controlled system reliability, serving as a proxy for subjective measurement. This approach treats trust as a static variable determined by system capability, overlooking individual cognitive differences and the dynamic trust calibration that naturally occurs across repeated takeover events. Since real-world trust is an evolving process, future work should integrate real-time measurements (e.g., psychophysiological monitoring) to capture the temporal dynamics of trust evolution.

Finally, constraints regarding the sample demographic (primarily young adults) and the driving simulator (lack of full motion cues) must be noted. Future studies should expand the participant pool to include broader age groups and validate findings through field operational tests to ensure ecological validity.

6. Conclusion

This study investigated the interactive effects of takeover warning strategies (Single-stage vs. Two-stage) and driver trust in automation (High vs. Low) on takeover behavior in Level 3 conditional automated driving through a human-in-the-loop driving simulation. The findings reveal a significant interaction between warning design and driver cognitive states.

First, the two-stage warning strategy demonstrated significant temporal benefits. Regardless of the driver's trust level, the provision of a preliminary MR effectively served as a cognitive prime, significantly shortening RTs. Second, trust level proved to be a determinant of driver vigilance. High trust induced automation complacency, resulting in prolonged reaction times, thereby confirming the inhibitory effect of over-trust on monitoring behavior.

Crucially, the study reveals that the safety benefits of two-stage warnings are not entirely robust against variations in trust levels. While the two-stage strategy optimized reaction speed, a significant interaction effect was observed in collision avoidance outcomes: high trust significantly undermined the intended safety advantages of the graded warning design, resulting in poorer avoidance performance compared to low-trust drivers. This suggests that optimizing warning architecture alone is insufficient to fully counteract the safety risks associated with trust-induced complacency.

In summary, this research highlights the necessity of jointly considering system characteristics and human cognitive states in takeover warning systems design. Effective takeover warning system should not only focus on temporal optimization but also integrate real-time trust monitoring and calibration mechanisms to ensure adequate performance even when drivers exhibit complacency. Future work will aim to validate these findings in more complex dynamic traffic environments and explore trust-aware adaptive warning strategies.

References

Agrawal, Shubham, and Srinivas Peeta (2021). "Evaluating the Impacts of Driver's Pre-Warning Cognitive State on Takeover Performance under Conditional Automation." *Transportation Research Part F: Traffic Psychology and Behaviour* 83 (November): 80–98.

Borowsky, Avinoam, Noa Zangi, and Tal Oron-Gilad (2022). "Interruption Management in the Context of Take-Over-Requests in Conditional Driving Automation." *IEEE Transactions on Human-Machine Systems* 52 (5): 1015–24.

Chen, Haolin, Xiaohua Zhao, Chen Chen, Zhenlong Li, Haijian Li, and Qiuhong Wang (2025). "A Systematic Review on Test Performance of the Driver Takeover Process in Automated Driving." *Accident Analysis & Prevention* 215 (June): 108012.

Chen, Haolin, Xiaohua Zhao, Haijian Li, Jianguo Gong, and Qiang Fu (2024). "Predicting Driver's Takeover Time Based on Individual Characteristics, External Environment, and Situation Awareness." *Accident Analysis & Prevention* 203 (August): 107601.

Deng, Hanwen, Guoliang Xiang, Jiandong Pan, et al (2024). "How to Design Driver Takeover Request in Real-World Scenarios: A Systematic Review." *Transportation Research Part F: Traffic Psychology and Behaviour* 104 (July): 411–32.

Gold, Christian, Riender Happee, and Klaus Bengler (2018). "Modeling Take-over Performance in Level 3 Conditionally Automated Vehicles." *Accident Analysis & Prevention*, Simulation of Traffic Safety in the Era of Advances in Technologies, vol. 116 (July): 3–13.

Hu, Chuan, Siwei Huang, Yu Zhou, et al (2023). "Dynamic and Quantitative Trust Modeling and Realtime Estimation in Human-Machine Co-Driving Process." SSRN Scholarly Paper No. 4647780. Rochester, NY, November 29.

Hu, Wenbo, Tingru Zhang, Yan Zhang, and Alan Hoi Shou Chan (2024). "Non-Driving-Related Tasks and Drivers' Takeover Time: A Meta-Analysis." *Transportation Research Part F: Traffic Psychology and Behaviour* 103 (May): 623–37.

Jian, Jiun-Yin, Ann M. Bisantz, and Colin G. Drury (2000). "Foundations for an Empirically Determined Scale of Trust in Automated Systems." *International Journal of Cognitive Ergonomics* 4 (1): 53–71.

Lu, Z., B. Zhang, A. Feldhütter, R. Happee, M. Martens, and J. C. F. De Winter (2019). "Beyond Mere Take-over Requests: The Effects of Monitoring Requests on Driver Attention, Take-over Performance, and Acceptance." *Transportation Research Part F: Traffic Psychology and Behaviour* 63 (May): 22–37.

Papadimitriou, Eleonora, Chantal Schneider, Juan Aguinaga Tello, Wouter Damen, Max Lomba Vrouenraets, and Annebel ten Broeke (2020). "Transport Safety and Human Factors in the Era of Automation: What Can Transport Modes Learn from Each Other?" *Accident Analysis & Prevention* 144 (September): 105656.

Petersen, Luke, Lionel Robert, X. Jessie Yang, and Dawn M. Tilbury (2019). "Situational Awareness, Drivers Trust in Automated Driving Systems and Secondary Task Performance." arXiv:1903.05251. Preprint, arXiv, March 12.

Qu, Jianhong, Zhou ,Ronggang, Zhang ,Yaping, and Qianli and Ma (2023). "Understanding Trust Calibration in Automated Driving: The Effect of Time, Personality, and System Warning Design." *Ergonomics* 66 (12): 2165–81.

Shahini, Farzaneh, and Maryam Zahabi (2022). "Effects of Levels of Automation and Non-Driving Related Tasks on Driver Performance and Workload: A Review of Literature and Meta-Analysis." *Applied Ergonomics* 104 (October): 103824.

Tan, Xiaomei, and Yiqi Zhang (2025a). "Driver Situation Awareness for Regaining Control from Conditionally Automated Vehicles: A Systematic Review of Empirical Studies." *Human Factors* 67 (4): 367–403.

Tan, Xiaomei, and Yiqi Zhang (2025b). "Understanding Driver Response to Multi-Stage Takeover Requests across Varied Modalities: A Computational Cognitive Modeling Approach." *Transportation Research Part C: Emerging Technologies* 174 (May): 105114.

Xu, Linli, Lie Guo, Pingshu Ge, and Xu Wang (2022). "Effect of Multiple Monitoring Requests on Vigilance and Readiness by Measuring Eye Movement and Takeover Performance." *Transportation Research Part F: Traffic Psychology and Behaviour* 91 (November): 179–90.

Yin, Weiru, Chen Chai, Ziyao Zhou, Chenhao Li, Yali Lu, and Xiupeng Shi (2021). "Effects of Trust in Human-Automation Shared Control: A Human-in-the-Loop Driving Simulation Study." *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, September, 1147–54.