Collision Reduction is Not the Whole Story: Injury-Oriented Safety Evaluation of Human–Vehicle Interaction Across Automated Driving Levels

Detong Qin, Jiajie Shen, Hanshuo Wang, Huili Yu, Kaiqin Zhao, Qing Zhou, Bingbing Nie

I. INTRODUCTION

Automated driving is widely expected to improve road traffic safety, yet comparable quantitative evidence on how safety performance changes across automation levels remains limited. Existing evaluations often focus on crash occurrence or takeover/intervention frequency [1][2], while injury consequences are not explicitly incorporated into cross-level safety assessment [3]. As a result, the protective value of automated driving may be incompletely characterised, especially when collisions are mitigated rather than fully avoided. Meanwhile, the driver's role changes fundamentally in automated vehicles, shifting from continuous vehicle control to supervisory intervention. When the automated system fails to fully resolve a hazard, the event enters a human–vehicle interaction regime in which driver intervention becomes a key determinant of the final outcome. Safety-critical events under automated driving should therefore be understood as outcomes jointly shaped by system capability and driver response, rather than by automation alone. Based on a comparable dataset of safety-critical human–vehicle interaction under automated driving, this study evaluates cross-level safety performance by jointly considering collision occurrence and injury consequences, while identifying the remaining interaction cases that warrant further attention.

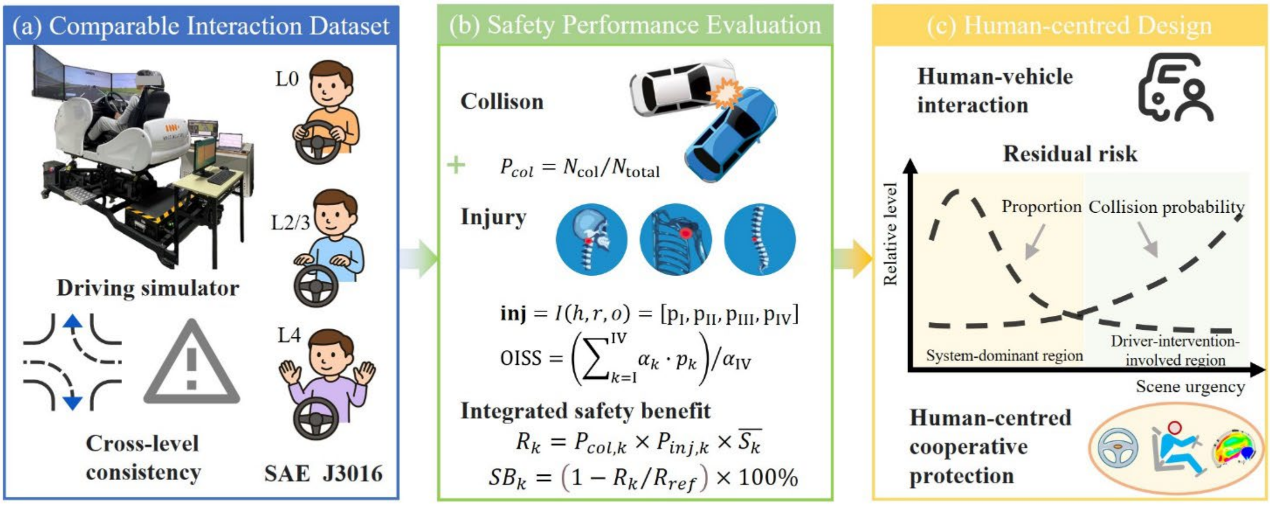

Fig. 1. Study framework for cross-level safety evaluation and human-centred cooperative protection in automated driving.

II. METHODS

The analysis was based on a driver-in-the-loop experimental platform covering manual driving (L0) and automated driving conditions (L2, L3, and L4), as defined in SAE J3016, under comparable experimental settings [4]. To enable cross-level comparison, the road type, safety-critical scenario set, and the distribution of scenario parameters were kept consistent across automation levels (Fig. 1(a)). A total of 2,582 safety-critical human–vehicle interaction cases were collected from 44 volunteers (30 males and 14 females). The dataset includes multiple types of data, including vehicle dynamics under automated control, the temporal characteristics of driver intervention (e.g., intervention latency and response intensity), and crash-case information. Together, these data provide a foundation for multi-dimensional analyses of safety performance and interaction in automated driving.

Three layers of outcome measures were emphasised in this study. First, the collision rate was used to describe the basic safety performance, including the human baseline collision rate, the system baseline collision rate, and the human–vehicle interaction collision rate. Second, post-crash injury severity was quantified to capture the consequence dimension beyond crash occurrence. Specifically, an RNN-based occupant injury prediction model was used to estimate the Occupant Injury Severity Score (OISS) [5], and the proportion of injury cases was

B. Nie (e-mail: nbb@tsinghua.edu.cn; tel: +86-10-6278-8689) is an Associate Professor, and Q. Zhou is a Professor in the School of Vehicle and Mobility at Tsinghua University, China. D. Qin, J. Shen and H.Wang are PhD students at the State Key Laboratory of Intelligent Vehicle Safety Technology and the School of Vehicle and Mobility, Tsinghua University, China. H. Yu is the Deputy Director of the State Key Laboratory of Intelligent Vehicle Safety Technology, China. K. Zhao is an engineer at Chongqing Changan Automobile Co., Ltd., China.

calculated to characterise post-crash injury risk. An integrated consequence-risk indicator was defined by combining collision rate, injury probability and injury severity, and was further transformed into an integrated safety-benefit indicator (Fig. 1(b)). In addition, a two-dimensional outcome matrix was defined based on intervention status (intervention / no intervention) and crash outcome (crash / no crash), in order to analyse the distribution and cross-level trends of interaction outcomes in automated driving.

III. INITIAL FINDINGS

Overall, in safety-critical scenarios, higher automation capability was associated with lower collision involvement, indicating a clear improvement in crash-avoidance performance. As automation capability increased, both the system baseline collision rate and the human–vehicle interaction collision rate decreased (Fig. 2(a)), with the latter declining from 25.19% to 15.37%. Notably, the human–vehicle interaction collision rate consistently lay between the system baseline and the human baseline, suggesting that the final safety outcome in automated driving reflects the joint effect of system capability and driver response.

However, protection-oriented evaluation requires attention not only to whether a crash occurs, but also to the resulting occupant injury. Injury-based results showed that although the probability of injury occurrence decreased, some crashes under higher levels of automated driving were associated with greater injury severity. Under the integrated safety-benefit indicator, L4 showed the best overall safety performance. By contrast, although L3 had a lower collision rate than L2, its overall safety benefit was reduced by a higher average injury severity (Fig. 2(b)).

Taken together, these findings suggest a safety performance pattern in which automated driving reduces overall collision involvement, while the remaining crash cases with elevated injury risk still deserve focused attention and targeted optimisation. From the perspective of human–vehicle interaction outcomes, the proportion of cases requiring driver intervention decreased with increasing automation level, yet crashes were concentrated in intervention-involved cases. Among all intervention cases, the conditional collision rate remained relatively stable at approximately 26% across automation levels (Fig. 2(c)). This indicates that although higher automation reduces overall collision risk, the remaining cases that still evolve into driver-intervention scenarios retain a similar likelihood of collision. The practically important extreme-risk cases are therefore likely concentrated in this residual subset of interaction cases, which deserve particular attention in follow-up work.

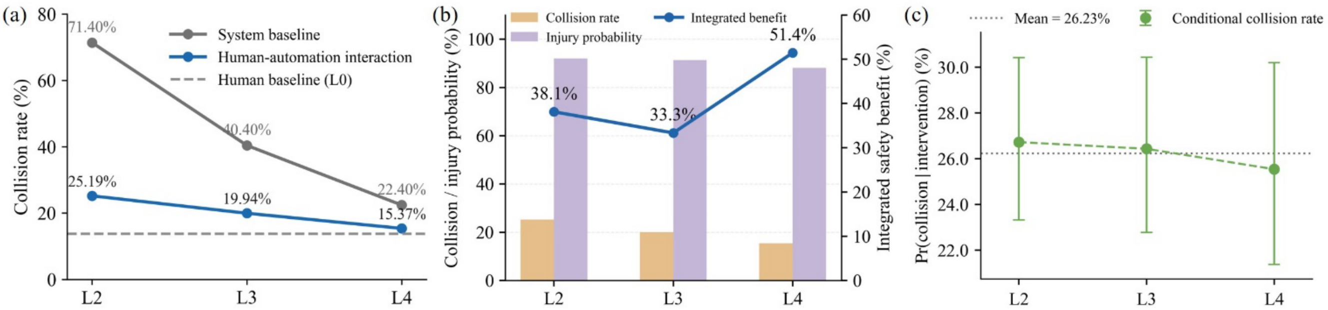

Fig. 2. Cross-level collision, injury, integrated safety benefit, and residual risk in human–vehicle interaction.

IV. DISCUSSION

The present findings indicate that automated driving safety does not improve linearly with automation level when evaluated in an integrated manner. Although collision involvement decreased with increasing automation capability, the corresponding safety benefit did not always improve monotonically once post-crash consequences were incorporated. This suggests that cross-level safety differences cannot be fully interpreted from crash avoidance alone. The approximately 26% intervention-conditioned collision rate further indicates that, across automation levels, a subset of safety-critical scenarios still evolves into collisions even after driver intervention. These remaining cases therefore deserve particular attention in efforts to improve occupant protection under automated driving. At the same time, the cross-level consistency of this pattern points to potentially common mechanisms in human–vehicle interaction, which deserve further investigation to clarify the formation of residual risk in automated driving. Future research may build on these findings to support more human-centred interaction design and protection-oriented cooperative safety strategies in automated vehicles (Fig. 1(c)).

V. ACKNOWLEDGEMENTS

This study is supported by the National Natural Science Foundation of China (52572413).

VI. REFERENCES

[1] Euro NCAP, Assisted Driving Test and Assessment Protocol, 2024.
[2] Wu, H., *et al., Sensors*, 2025.
[3] Ding, S., *et al., Accid. Anal. Prev.*, 2024.
[4] Shen, J., *et al*., *Accid Anal Prev*,2025.
[5] Wang, Q., *et al*., *iScience*, 2022.