**Modeling takeover decisions in driving automation: A multilevel drift–diffusion model (MDDM) framework integrating human, system, and environmental factors**

Youyu Sheng

State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100000, China

The Hong Kong University of Science and Technology (Guangzhou), Guangdong 510000, China

Email: ysheng330@connect.hkust-gz.edu.cn.

Chouyu Wu

School of Computer Science and Engineering, Sun Yat-sen University, Guangdong, China

Email: wuchy66@mail2.sysu.edu.cn

Songmei Li

The Hong Kong University of Science and Technology (Guangzhou), Guangdong 510000, China

Email: sli861@connect.hkust-gz.edu.cn

Yufan Chen

The Hong Kong University of Science and Technology (Guangzhou), Guangdong 510000, China

Email: ychen729@connect.hkust-gz.edu.cn

Xianghong Sun

State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100000, China.

Department of Psychology, University of Chinese Academy of Sciences, Beijing 100000, China.

Email: sunxh@psych.ac.cn

Dengbo He

Thrust of Robotics and Autonomous Systems, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511400, China.

Thrust of Intelligent Transportation, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511400, China.

Email: dengbohe@hkust-gz.edu.cn

Jingyu Zhang

State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100000, China.

Department of Psychology, University of Chinese Academy of Sciences, Beijing 100000, China.

Email: zhangjingyu@psych.ac.cn

**Corresponding author.** Dengbo He; Jingyu Zhang

**Abstract**

Takeover behaviors represent one of the most critical safety challenges when driving with automation, yet existing models for takeover prediction often overlook the underlying cognitive mechanisms. To address this gap, we propose a multilevel modeling framework that integrates hierarchical linear modeling with a drift–diffusion model (DDM) to characterize how human, system, and environmental factors jointly shape the cognitive process of takeover decisions. The framework separates rapid fluctuations in environmental influences from stable human and system influences and maps these components onto distinct cognitive parameters. To evaluate the framework, we collected a driving-simulator dataset (N = 64) that systematically manipulated vehicle speed, optical flow, lighting, driving experience, and automation reliability. The fitted model reproduced response-time distributions that highly matched empirical data, indicating high psychological interpretability and modeling validity. These results revealed a dissociable cognitive structure: environmental factors accelerated evidence accumulation by increasing drift rates; experienced drivers adopted higher and more stable decision boundaries; and lower perceived reliability reduced non-decision time by increasing monitoring engagement. These findings demonstrate that takeover behavior emerges from distinct cognitive pathways. Moreover, the proposed framework offers a generalizable framework to integrate human, environmental, and system factors into cognitive process models in the context of human–intelligent system interaction.

**Keywords: Driving automation; Takeover decisions; Drift–diffusion model; Hierarchical linear modeling; Risk perception**

## 1. Introduction

Driving automation technologies are rapidly advancing, yet ensuring safety during transitions of control remains a critical challenge (Bathla et al., 2022; L. Chen et al., 2024; Hasegawa et al., 2024; Miqdady et al., 2023). A growing number of crashes and near-miss incidents have highlighted that takeover behavior poses a significant safety challenge in human–automation interaction (Barisic et al., 2023; Miller et al., 2024; Morales-Alvarez et al., 2020). In such situations, drivers must quickly perceive dynamic risks, evaluate the reliability of system cues, and execute timely control actions. The prediction of takeover decisions, therefore, plays a decisive role in the overall safety and public acceptance of automated vehicles (Soares et al., 2021; Tan & Zhang, 2022; Wu et al., 2022).

Previous research predicting takeover decisions has employed a wide range of modeling approaches, such as classical machine learning models (Ayoub et al., 2022; H. Chen et al., 2024; Du et al., 2020; Gipps, 1981; Holley et al., 2023; Liu et al., 2022; Pakdamanian et al., 2021; Treiber et al., 2000; C. Wang et al., 2025). Despite substantial progress, most existing behavioral models still oversimplify the cognitive processes that underlie drivers' interventions. These models typically treat takeover as a stimulus–response mapping, without decomposing the risk-based decision process underlying drivers' interventions. More specifically, takeover in automated driving is not merely a reflexive reaction to a warning signal. When a takeover request occurs, drivers must first encode risk cues from the driving scene, accumulate risk-related evidence from dynamic environmental information, evaluate whether the accumulated evidence is sufficient to justify intervention, and translate this decision into manual control. Existing models rarely distinguish among these process components. As a result, they cannot determine whether a delayed takeover reflects weak risk-cue encoding, slow accumulation of risk-related evidence, a conservative intervention threshold, or reduced perceptual–motor readiness during supervisory control. This oversimplification presents a critical barrier to both model performance and practical application. First, due to ignorance of the temporal dynamics of evidence accumulation, these models lack the robustness to generalize across complex, dynamic driving environments. Second, the lack of psychological explainability renders these models as 'black boxes'—they can predict when a takeover occurs, but not why it is delayed. Crucially, this opacity limits the model's ability to guide downstream interventions. Without distinguishing whether a performance deficit stems from slow perceptual encoding (e.g., needing better visual cues) or excessive response caution (e.g., needing confidence training), it is infeasible to design adaptive support systems or targeted driver training programs that effectively address the root cause of the behavior.

Therefore, recent studies have begun to introduce the drift–diffusion model (DDM) into driving research to describe the temporal dynamics of decision-making (Durrani & Lee, 2023; Hu et al., 2025; Huang et al., 2025; Myers et al., 2022). The DDM has been widely used in cognitive psychology to explain how individuals accumulate uncertain information over time before committing to an action (Ratcliff et al., 2016). Its relevance to takeover modeling lies in its ability to decompose response time into theoretically meaningful cognitive processes underlying takeover timing. Specifically, drift rate captures the speed and quality of risk-related evidence accumulation; decision boundary captures the amount of evidence required before committing to a takeover action, reflecting response caution or intervention threshold; and non-decision time captures perceptual encoding and motor execution processes outside evidence accumulation. In this way, the DDM allows takeover time to be interpreted not merely as a faster-or-slower behavioral response, but as the joint outcome of evidence accumulation, intervention-threshold setting, and perceptual–motor readiness. In addition, because the DDM models the distribution of response times rather than only mean response latency, it provides a useful process-level framework for examining how different cognitive processes jointly shape takeover timing.

Despite these advantages, existing DDM-based driving studies have several limitations when applied to takeover decisions in automated driving. First, most prior applications have focused on manual driving contexts, such as car-following, braking, or high-risk maneuver decisions, in which drivers are already engaged in continuous vehicle control (Durrani & Lee, 2023; Huang et al., 2025). Automated takeover differs from these contexts because drivers must transition from supervisory monitoring to manual re-engagement under uncertainty and time pressure. This transition makes perceptual–motor readiness and re-entry into the control loop particularly important.

Second, existing DDM applications often emphasize vehicle kinematic cues, such as speed or time-to-collision, and explain behavioral variability primarily through changes in drift rate. As a result, other cognitive processes involved in takeover decisions have received far less attention, including intervention-threshold setting and non-decision processes related to perceptual encoding and motor preparation.

Third, prior approaches rarely account for the multilevel nature of automated takeover decisions. Environmental factors such as vehicle speed, optical flow, and lighting vary across trials and shape moment-to-moment risk evidence, whereas driving experience and automation reliability represent relatively stable human and system-level influences. Without separating these levels of influence, existing DDM frameworks may conflate transient environmental effects with stable individual or system-related factors. These limitations make it difficult to directly apply existing DDM frameworks to the unique risk-based decision processes involved in takeover behavior under driving automation.

To address these limitations, the present study applies the DDM to capture the cognitive mechanisms underlying takeover decisions while integrating environmental, system, and human factors into the modeling framework. This approach provides a more comprehensive and quantitative account of how drivers make takeover decisions.

## 2. Related work

### 2.1 Machine Learning Models for Takeover Prediction

Previous work in general employed machine learning (ML) approaches to capture complex, nonlinear relationships between multimodal inputs and driver responses and thus predict takeover performance. These models typically integrate vehicle dynamics, driver state indicators, and environmental features, and in some cases incorporate physiological or eye-tracking data, to predict outcomes such as reaction time, takeover readiness, or takeover quality.

For example, Du et al. (2020) evaluated six supervised learning algorithms using physiological and situational features collected prior to takeover requests, finding that a random forest classifier yielded the most reliable prediction of takeover performance. Pakdamanian et al. (2021) proposed a deep neural network that fuses vehicle, physiological, and subjective measures to jointly predict takeover intention, time, and quality with high accuracy. Similarly, Ayoub et al. (2022) applied XGBoost combined with Shapley additive explanations (SHAP) to identify key predictors of takeover time from a large meta-analytic dataset. H. Chen et al. (2024) further extended this approach by integrating individual difference variables and situational awareness metrics to improve generalizability.

Collectively, these models have substantially improved predictive accuracy in takeover decision prediction. However, current ML models remain black boxes; their explainability to uncover why and how certain variables influence human decision-making remains weak. Even with post-hoc explainability methods such as SHAP, the insights gained are statistical rather than psychological: they reveal which features matter most but not the underlying cognitive rationale.

### 2.2 Drift–Diffusion Model for Takeover Prediction

The DDM has been widely used in cognitive psychology and decision sciences to describe how individuals integrate uncertain information over time before committing to an action (Hu et al., 2025; Huang et al., 2025; Myers et al., 2022; Ratcliff et al., 2016). Recent work has begun to apply DDMs to manual driving research. For instance, Durrani and Lee (2023) developed an accumulator model in which drivers' reaction times in car-following were predicted from visual looming information—quantified as the inverse time-to-contact—that accumulates until a decision threshold is reached. Similarly, Huang et al. (2025) introduced a risk-sensitive DDM that combines kinematic indicators with clustering methods to characterize drivers' risk sensitivity and map these profiles onto decision parameters. These studies demonstrated that DDM-based approaches can capture the temporal dynamics of evidence accumulation in emergency maneuvers more effectively than traditional car-following or lane-changing models.

Nevertheless, the generalizability of these findings is limited, as the continuous control inherent in manual driving differs fundamentally from the abrupt re-engagement required in vehicles with driving automation. As a result, these models do not fully capture the distinct cognitive and temporal dynamics of takeover decisions in vehicles with driving automation. In addition, existing models rarely incorporate environmental factors or human factors in a systematic manner, limiting their ability to capture the ecological complexity inherent in takeover scenarios. More fundamentally, in prior studies, behavioral variability is assumed to arise primarily from changes in the drift rate. This drift-centric perspective overlooks other theoretically important components of the decision process, such as decision boundary setting and non-decision time, which capture response caution and perceptual–motor delays, respectively. Therefore, existing DDM approaches provide an incomplete account of how takeover decisions unfold and limit insight into the cognitive processes underlying human–automation transitions.

### 2.3 The Current Study

Building upon these limitations, the current study aims to apply the DDM to takeover behavior when driving with automation and further examines how to integrate multiple classes of factors (human, system, and environment) into the DDM framework. However, directly incorporating multiple factors into a DDM framework poses several challenges. First, human, system, and environmental factors differ fundamentally. Human traits such as driving experience represent stable individual dispositions that remain relatively constant across situations. Similarly, system properties such as automation reliability are typically fixed within a driving session and do not fluctuate in real time. In contrast, environmental factors such as speed or illumination fluctuate continuously and shape drivers' moment-to-moment risk perception. Treating all these influences as if they varied on the same timescale risks conflating stable dispositions with transient situational dynamics, thereby reducing psychological interpretability and ecological validity.

Second, beyond temporal differences, there is the critical challenge of individual heterogeneity. The influence of these human, system, and environmental factors is rarely uniform across the population. Instead, drivers exhibit substantial variability in their sensitivity to these constraints—for instance, the same level of automation unreliability may trigger rapid intervention in one driver while causing hesitation in another. Traditional approaches that assume fixed effects fail to capture these driver-specific, distinct sensitivities, potentially leading to biased estimations of the decision-making process. These challenges collectively constrain the application of DDM in takeover research.

To address these challenges, we propose a two-stage framework that integrates hierarchical linear modeling (HLM) with the DDM to model takeover decisions in the context of driving automation. In the first stage, HLM separates trial-level environmental variation from relatively stable human and system-level influences. Environmental factors such as vehicle speed, optical flow, and lighting vary across trials and shape moment-to-moment risk-related evidence, whereas driving experience and automation reliability represent more stable individual and system-level factors. This hierarchical separation reduces the risk of conflating transient environmental effects with stable driver- or system-related influences and provides structured inputs for the subsequent decision model.

In the second stage, the resulting factor classes are mapped onto distinct DDM components to represent multiple cognitive processes involved in takeover decisions. Environmental factors are linked to drift rate because they shape the speed and quality of risk-related evidence accumulation (Buzon et al., 2021; Köhler et al., 2023; Köhler et al., 2022; Wood, 2020; Zhou et al., 2024). Driving experience is linked to decision boundary because it may influence intervention-threshold setting, that is, the amount of evidence required before committing to a takeover action (He et al., 2021; Sheykhfard et al., 2022; Sun & He, 2025). Automation reliability is linked to non-decision time because it may affect monitoring engagement and perceptual–motor readiness during supervisory control (Beller et al., 2013; Li et al., 2025; Schwarz et al., 2019).

By integrating these two stages, the proposed MDDM contributes a process-level modeling framework for automated takeover decisions. Rather than treating takeover behavior as a single-factor or drift-centric process, the framework jointly incorporates heterogeneous environmental, human, and system-level determinants and distributes their effects across multiple theoretically meaningful cognitive processes, including risk-related evidence accumulation, intervention-threshold setting, and perceptual–motor readiness. In doing so, the framework provides a principled bridge between environmental conditions, driver characteristics, automation reliability, and the cognitive processes through which they shape risk-based takeover timing.

## 3. Data Collection

Implementing this MDDM framework requires a dataset that captures environmental, human, system factors, and takeover behavior within a unified structure. Existing datasets do not provide this level of integration, as they typically focus on either vehicle kinematics or isolated behavioral measures. Therefore, the present study begins by constructing a dedicated takeover dataset with various human, driving automation system and environment conditions, which provides the necessary empirical foundation for modeling.

### 3.1 Selection of Experimental Factors

Before collecting data, we first identified the most representative variables at the human, system, and environmental factors to serve as the independent variables in the study. At the human level, driving experience was selected because prior research consistently demonstrates that experience strongly shapes hazard anticipation, situation assessment, and the ability to take over control during automation failures (He et al., 2021; Sun & He, 2025). At the system level, automation reliability was chosen, given its well-established role in influencing trust calibration, perceived system predictability, and drivers' intervention thresholds (Lee & Pitts, 2024; Li et al., 2025; Schwarz et al., 2019).

Prior research has identified a wide range of environmental factors that influence drivers' risk perception and takeover behavior. Among these, optical flow, lighting of the scenarios, and vehicle speed are the most frequently examined and experimentally tractable dimensions in automated-driving studies. Optical flow reflects the rate of visual expansion and provides critical information for time-to-collision judgments (Köhler et al., 2023; Köhler et al., 2022, 2025); lighting determines the clarity and reliability of visual information in the driving scene (Hu et al., 2022; Wood, 2020; Zhou et al., 2024); and vehicle speed directly modulates the kinematic urgency associated with an approaching hazard (Buzon et al., 2021; Ju & Wallraven, 2023; Mohammad et al., 2023). Accordingly, optical flow, lighting, and vehicle speed were selected as representative environmental variables in our experiment.

### 3.2 Participant

We recruited 64 licensed drivers from China, comprising 32 males and 32 females, with an average age of 25.58 years (SD = 4.30). Participants had to possess a valid driver's license, be righthanded, have normal or corrected-to-normal vision, and have no history of neurological issues or color-vision deficiencies. Two distinct groups of participants were recruited based on their years of licensure and the distance driven in the previous year. Experienced drivers needed to have held their licenses for 3 or more years and driven over 5,000 km in the past year, while novice drivers were required to have had their licenses for less than a year and driven under 5,000 km in the same period. Approval for the study was granted by the Ethics Committee of [Placeholder for double-blinded review].

**Table 1. Demographic information of participants.**

| **Characteristic** | **Low experience group (n=32)** | | **High-experience group (n=32)** | |
|---|---|---|---|---|
| **Reliability** | Low reliability group | High reliability group | Low reliability group | High reliability group |
| **Gender** | | | | |
| Male | 8 | 8 | 8 | 8 |
| Female | 8 | 8 | 8 | 8 |
| **Age** | | | | |
| Mean | 24.19 | 22.44 | 26.94 | 28.75 |
| SD | 3.75 | 2.39 | 3.57 | 4.45 |
| [Min, Max] | [19, 35] | [19, 25] | [21, 35] | [24, 37] |

| | | | | |
|---|---|---|---|---|
| **Driving year** | <1 year | <1 year | ≥ 3 years | ≥ 3 years |
| **Annual mileage** | <5,000 km | <5,000 km | ≥5,000 km | ≥5,000 km |

## 3.3 Experiment Design

The experiment adopted a mixed factorial design consisting of two between-subject factors and three within-subject factors. The between-subjects factors include System Reliability (high vs. low) and Driving Experience (high vs. low). Experienced and novice drivers were randomly assigned to either the high or low reliability condition, resulting in four experimental groups, each with 16 participants. Each participant completed eight takeover trials that fully crossed three repeated-measures factors: Optical Flow (high vs. low), Lighting Condition (daytime vs. nighttime), and Vehicle Speed (high vs. low). To minimize carryover and learning effects, the presentation order of these eight experimental scenarios was counterbalanced across participants using a Latin Square design. As a result, the overall study design constituted a 2 × 2 (between-subjects) × 2 × 2 × 2 (within-subject) mixed design.

## 3.4 Manipulation of Independent Variables

### *3.4.1 System Reliability*

To influence participants' perception of system reliability, we followed an established manipulation approach used in driving automation research (Youyu Sheng et al., 2025). All participants first read a brief, neutral introduction emphasizing that the driving automation required continuous human supervision and occasional intervention. They were then provided with a standardized reliability briefing describing the past test performance and independent ranking of the assigned system brand. Two versions of this briefing were developed:

**High-reliability**: The brand assigned (WLE) was portrayed as having undergone 1,000 km of testing in various road and weather conditions, needing just two human interventions. It was also noted to have ranked a 2nd place in an independent evaluation of mainstream driving automation systems.

**Low-reliability**: The assigned brand (DID) underwent testing over the same 1,000 km under identical conditions but required nine driver interventions. In the independent evaluation, it was ranked 14th.

Participants were told that the simulated driving automation system is from the respective brand. However, in both scenarios, the driving situations and simulator control system were kept the same, the only difference was in the content of the system reliability briefing. In addition, the brand labels in this experiment were entirely fictitious, with no actual market recognition or brand background, thus eliminating the confounding effects of brand-related factors on the manipulation of system reliability.

### *3.4.2 Environmental Variables*

***Lighting condition*** was manipulated by presenting scenarios at two times of day: daytime and nighttime. In the daytime condition, the driving environment was illuminated by natural daylight, whereas in the nighttime condition, ambient illumination was substantially reduced, and roadway visibility relied primarily on artificial lighting.

***Vehicle speed*** during driving was manipulated at two levels: 40 km/h (low speed) and 100 km/h (high speed). These speeds were selected to represent typical urban and highway conditions to systematically vary the kinematic risk associated with the takeover event.

***Optical flow*** was manipulated by varying the spacing of roadside streetlights, following established approaches in driving perception research (Köhler et al., 2025), with inter-lamp distances set to 10 m (high optical flow) and 30 m (low optical flow). Shorter spacing increased the density and temporal frequency of visual motion cues, whereas longer spacing reduced visual flow information.

### 3.5 Measurement

***P*erceived risk (PR).** The PR was measured using the Multidimensional Perceived Risk Scale (MPRS) to evaluate the effectiveness of the experiment control (Y. Sheng et al., 2025). Participants rated items on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). The scale has demonstrated strong internal consistency (α > .90) in driving automation systems. The final score in this study was calculated as the mean value of all 27 items in the scale. Notably, the 3 items of the Performance Integrity Risk (PIR) dimension were reverse-coded, and all these items were recoded before score calculation, while all items in other dimensions were positively coded. After score computation, the final PR scores were standardized for analysis to facilitate comparability across participants, with higher standardized values indicating a higher level of perceived risk among drivers.

***Takeover time***. Takeover time was defined as the elapsed time between the onset of a takeover request (TOR) and the driver's first effective control maneuver, operationalized as either a steering wheel angle change exceeding 2° or a brake pedal actuation exceeding 10% (Gold et al., 2013). The takeover time is one of the most widely adopted measures of takeover performance, and it was found to be associated with crash risk in takeover events (Soares et al., 2021).

### 3.6 Apparatus

The experiment was conducted using a fixed-base driving simulator (Info Tech, Shanghai; see Figure 1a). Driving scenarios were created with SILAB 7.1 software and displayed on three 43-inch screens situated about 1 meter from the driver. Each screen had a resolution of 1920 × 1080 pixels, providing a horizontal field of view of 150° and a vertical field of view of 47°, creating an immersive driving experience. Participants could toggle the driving automation system (ADS) on or off using a virtual button on a 15.6-inch interactive touchscreen located next to the steering wheel. Manual control could be regained by turning the steering wheel at least 2° or by pressing the brake pedal with a minimum force of 10% (Gold et al., 2013). Driving performance data were captured at a frequency of 60 Hz with SILAB 7.1 software, and all data streams were synchronized using the Human Research Tool (HRT) software from Info Tech.

### 3.7 Driving Task

The driving task involved a simulated highway setting with eight scenarios. In each scenario, participants initiated the session by activating the ADS. After around three minutes of driving with automation, they approached a simulated road construction site (see Figure 1c). A TOR was triggered when the vehicle was 150 meters from the construction area (Banks & Stanton, 2016), though a different roadside environment was adopted in different trials. This TOR, communicated via an auditory warning tone, instructed drivers to resume manual control. If there was no response, the ADS would continue its path and eventually collide with the construction barrier. Participants were required to disengage the automation, take over, and safely maneuver through the construction zone (see Figure 1d).

## 3.8 Procedure

Upon arriving at the laboratory, participants were given the experimental instructions to review and had the opportunity to ask any questions. They were explicitly told that the driving automation system was not fully autonomous and that they were responsible for monitoring the vehicle and ensuring safety throughout the session. Afterward, participants entered the fixed-base driving simulator, viewed the introductory materials for their assigned condition, and then began the driving task.

The entire experimental session lasted approximately 90 minutes. Upon completing all driving scenarios, participants were debriefed, compensated, and thanked for their participation.

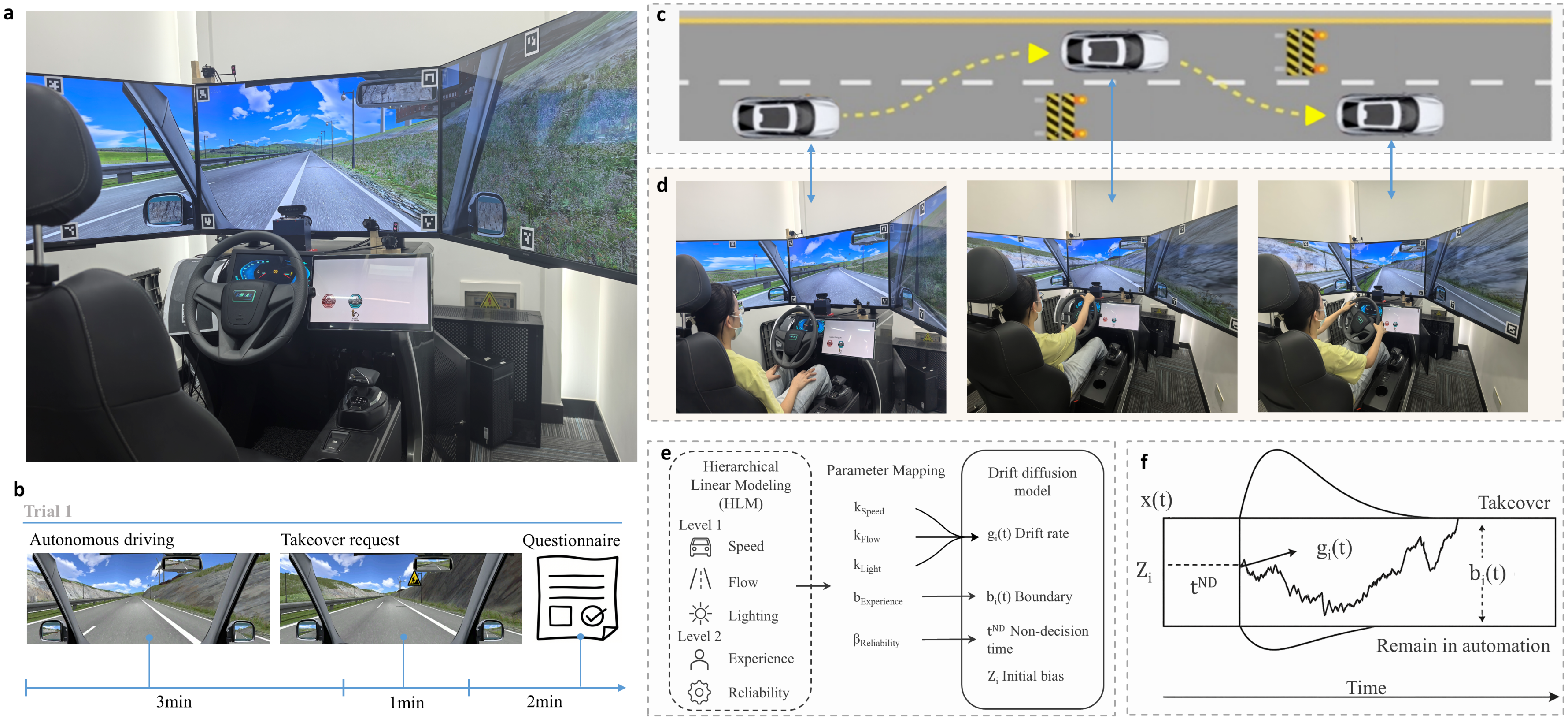


**Figure 1. Study overview. (a) Experimental apparatus. (b) Trial structure in the driving task. (c) Illustration of the takeover scenario. (d) Participant performing the takeover maneuver. (e) Schematic illustration of the modeling process. (f) Drift–diffusion process for takeover decision.**

## 4. MDDM Modeling Framework

Understanding takeover decisions requires a modeling framework that can jointly capture the distinct roles of environmental, human, and system-level factors while respecting the different temporal scales at which these factors operate. To achieve this, we implement a two-stage modeling framework that integrates HLM with a DDM. As illustrated in Figure 2, the experimental dataset contains both perceived risk data and behavioral data (takeover time) collected under the same manipulated conditions. In the first stage, perceived risk data are entered into the HLM to extract structured components associated with environmental, human, and system factors. In the second stage, these HLM-derived components are mapped onto distinct DDM parameters, which are then used to model takeover decision dynamics and predict takeover response times.

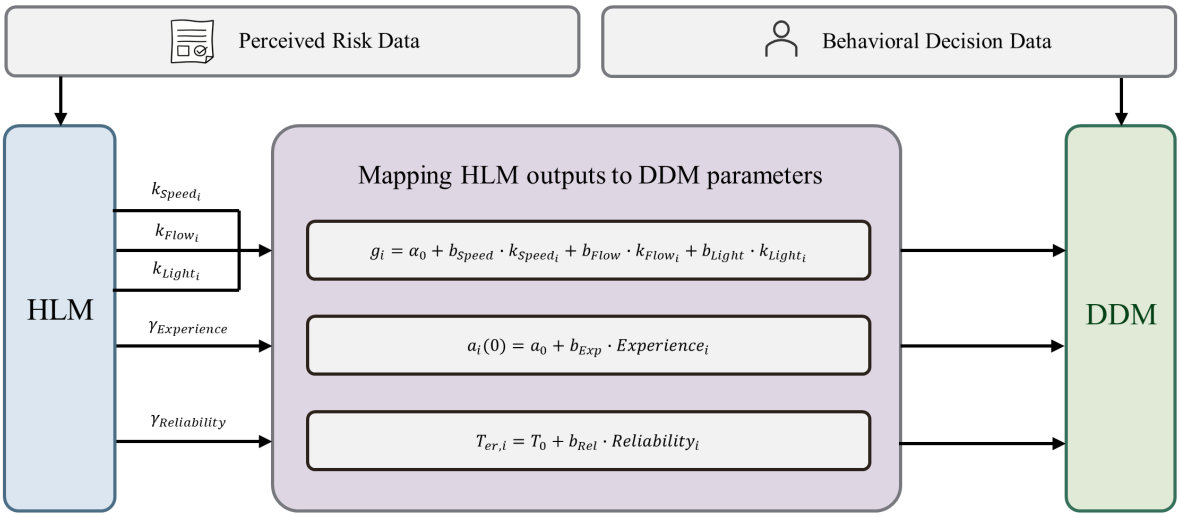


**Figure 2. Overview of the proposed MDDM modeling framework.**

### 4.1 Hierarchical Linear Modeling of Perceived Risk

Building on this two-stage framework, we first used an HLM to describe drivers' perceived risk as a joint function of Level-1 factors (speed, optical flow, and lighting) and Level-2 factors (driving experience and perceived system reliability). Before specifying the multilevel structure, we calculated the intraclass correlation coefficient (ICC) to quantify the extent of stable between-driver differences in perceived risk. The estimated ICC was 0.73, indicating that approximately 73% of the variance in perceived risk was attributable to between-driver variability. This substantial between-driver variance supports the inclusion of random effects in the multilevel model.

### *4.1.1 Level 1 Model: Environmental Factors*

At the Level 1 model, trial-level variation in perceived risk is modeled as a function of the three within-subject environmental factors manipulated during the experiment:

$$Y_{ij} = \beta_{0_i} + k_{\text{Speed}_i} \cdot \text{Speed}_{ij} + k_{\text{Flow}_i} \cdot \text{Flow}_{ij} + k_{\text{Light}_i} \cdot \text{Light}_{ij} + \varepsilon_{ij} \tag{1}$$

$\beta_{0_i}$ is represents driver i's baseline level of risk perception, $k_{\text{Speed}_i}$, $k_{\text{Flow}_i}$ and $k_{\text{Light}_i}$quantify how strongly each environmental factor contributes to the driver's perceived risk. Together, these parameters define the individual's risk perception structure. The trial-level residual $\varepsilon_{ij}$ captures trial-specific unexplained variance and is assumed to follow a normal distribution with mean of 0 and variance of $\sigma_\varepsilon^2$.

In addition, in the present HLM specification, the Level 1 environmental predictors were modeled as fixed slopes, whereas only the intercept was allowed to vary randomly across drivers. This specification was adopted to maintain model parsimony and estimation stability, given the limited number of repeated observations per participant and the primary aim of separating within-driver environmental effects from stable between-driver differences in baseline perceived risk.

#### *4.1.2 Level 2 Model: Human and System Factors*

At the Level 2 level, we modeled between-driver variation in the baseline parameter $\beta_{0i}$, which reflects each driver's overall level of perceived risk across situations. Specifically, we examined whether this baseline differed systematically as a function of the two between-subjects experimental factors in the study, i.e., driving experience and system reliability:

$$\beta_{0_i} = \beta_0 + \gamma_{\text{Experience}} \cdot \text{Experience}_i + \gamma_{\text{Reliability}} \cdot \text{Reliability}_i + u_{0i} \tag{2}$$

$\beta_0$ represents the grand mean baseline of perceived risk across all drivers. The fixed effects $\gamma_{\text{Experience}}$ and $\gamma_{\text{Reliability}}$ quantify how driving experience and system reliability shift a driver's baseline risk perception. The random effect $u_{0i}$ captures residual individual differences not explained by these predictors and is assumed to follow a normal distribution with variance $\sigma_{u0}^2$.

This hierarchical decomposition cleanly separates environmental, human, and system influences, yielding interpretable parameters while preserving their distinct temporal and psychological roles in the DDM.

#### *4.1.3 HLM Estimation Results*

To facilitate interpretation of the model coefficients, all predictors were effect-coded as −0.5 and +0.5. Specifically, high vehicle speed, high optical flow, daytime lighting, high system reliability, and high driving experience were coded as +0.5. Low vehicle speed, low optical flow, nighttime lighting, low system reliability, and low driving experience were coded as −0.5.

Table 2 presents the fixed-effect estimates. Driving experience showed a significant effect on perceived risk (b = −0.567, SE = 0.211, 95% CI [−0.989, −0.146], $p$ = .009). The effects of vehicle speed, optical flow, lighting condition, and system reliability were not statistically significant in this HLM summary model.

**Table 2. Fixed-Effect Estimates from the HLM Predicting Perceived Risk**

| Predictor | Estimate | SE | df | 95% CI | t | p |
|---|---|---|---|---|---|---|
| Intercept | −0.012 | 0.106 | 64.25 | [−0.222, 0.199] | −0.110 | .912 |
| Vehicle speed | 0.071 | 0.043 | 419.51 | [−0.014, 0.156] | 1.650 | .100 |
| Optical flow | −0.058 | 0.043 | 419.44 | [−0.143, 0.027] | −1.349 | .178 |
| Lighting condition | −0.040 | 0.043 | 420.38 | [−0.125, 0.045] | −0.920 | .358 |
| System reliability | −0.078 | 0.211 | 64.25 | [−0.499, 0.344] | −0.369 | .713 |
| Driving experience | −0.567 | 0.211 | 64.25 | [−0.989, −0.146] | −2.687 | .009 |

The random-intercept variance was 0.682, and the residual variance was 0.223. The model-fit indices were as follows: log-likelihood = −424.269, AIC = 864.538, BIC = 897.978. Compared with the null model, the full HLM showed significantly improved model fit, $\chi^2(5) = 12.367$, $p$ = .030.



## 4.2 Drift Diffusion Modeling of Takeover Decision

Building on the HLM estimates derived in Section 4.1, we next modeled how these perceptual and individual-difference components translate into drivers' takeover decisions using a DDM. The DDM provides a mechanistic account of how evidence is accumulated over time until a decision threshold is reached, and it allows the Level 1 environmental parameters and Level 2 human and system parameters obtained from the HLM to influence distinct components of the decision process in a theoretically coherent manner.

Therefore, we implemented a DDM in which the decision process is decomposed into three core components: the drift rate, the decision boundary, and the non-decision time. Each component is mapped to a distinct class of influencing factors identified in driving automation to preserve their effects. The following subsections detail the formulation of each component and its integration into the complete model.

### *4.2.1 Drift Rate*

The drift rate $g_i$ represents the average rate at which driver *i* accumulates decision evidence over time, determining the speed and direction of the takeover decision. Environmental factors were mapped to the drift rate because hazard perception theory suggests that visual and kinematic cues in the driving environment shape the overall strength of perceptual evidence available for danger detection and urgency appraisal (Horswill, 2016). Because drift rate represents the quality and direction of the sensory evidence that is accumulated over time, it must be determined by cues that directly shape overall strength of perceptual evidence available to the driver. In driving, environmental factors such as vehicle speed, optical flow, and lighting constitute the primary sources of perceptual evidence that inform hazard appraisal and urgency judgments (Buzon et al., 2021; Köhler et al., 2023; Köhler et al., 2022; Wood, 2020; Zhou et al., 2024). Accordingly, these environmental factors serve as the inputs to the drift rate, allowing the evidence accumulation process to reflect how environmental factors influence the perceived risk of an impending hazard.

These parameters $k_{\text{Speed}_i}$, $k_{\text{Flow}_i}$ and $k_{\text{Light}_i}$ are estimated from the HLM introduced in Section 4.1 and capture individual differences in how drivers appraise changes in the driving environment. Accordingly, the drift rate for driver *i* is modeled as a linear combination of these perceptual weights:

$$g_i = \alpha_0 + b_{\text{Speed}} \cdot k_{\text{Speed},i} + b_{\text{Flow}} \cdot k_{\text{Flow},i} + b_{\text{Light}} \cdot k_{\text{Light},i} \qquad (3)$$

Here, $\alpha_0$ denotes the baseline drift, and $b_{\text{Speed}}, b_{\text{Flow}}, b_{\text{Light}}$ are coefficients determining how the three environmental factors contribute to the drift rate. This formulation links individual differences in environmental risk perception directly to the evidence accumulation process.

#### *4.2.2 Decision Boundary*

The decision boundary $a_i(t)$ determines the amount of accumulated evidence required for driver *i* to initiate a takeover response. A higher boundary reflects greater response caution, whereas a lower boundary corresponds to faster but potentially less conservative decisions. Driving experience was mapped to the decision boundary based on naturalistic decision-making theory (Klein, 1993), particularly the recognition-primed decision account. This perspective suggests that, in dynamic and time-pressured environments, experience primarily supports more stable situational recognition and better calibrated action criteria, rather than simply faster extraction of sensory evidence. Since the decision boundary reflects the amount of evidence required before committing to action, driving experience was therefore considered more appropriately linked to boundary setting. Consistent with prior driving research, experienced drivers typically adopt more consistent and conservative criteria for initiating control actions (He et al., 2021; Sheykhfard et al., 2022; Sun & He, 2025). Accordingly, driving experience is incorporated into the initial decision boundary, allowing human factors to shape the threshold-setting process. To capture the increasing urgency present in hazardous situations, the boundary is modeled as an exponentially collapsing function.

First, each driver's initial boundary height is expressed as a function of driving experience:

$$a_i(0) = a_0 + b_{\text{Exp}} \cdot \text{Experience}_i, \qquad (4)$$

where $a_0$ represents the population-level baseline boundary, and $b_{\text{Exp}}$ captures how driving experience modulates the initial level of decision caution. The boundary then decreases over time according to:

$$a_i(t) = a_{\min} + (a_i(0) - a_{\min})e^{-kt}, \qquad (5)$$

where $a_{\min}$ is the minimum allowable threshold, and $k > 0$ controls the rate of collapse. A larger $k$ indicates a faster reduction in decision caution as the urgency of the situation increases. The inclusion of $a_{\min}$ prevents the threshold from collapsing to zero, ensuring numerical stability while reflecting realistic lower bounds on decision caution. This parameterization links individual differences in baseline caution to a psychologically interpretable, time-varying decision threshold and allows the model to represent how drivers relax their criteria as hazard urgency escalates.

#### *4.2.3 Non-decision Time*

Non-decision time $T_{\mathrm{er},i}$ captures the components of the takeover response that do not involve evidence accumulation, including perceptual encoding before accumulation begins and motor execution after a decision boundary is crossed. System reliability was mapped to non-decision time based on trust calibration theory (Lee & See, 2004). Prior theory suggests that perceived reliability affects drivers' monitoring engagement, vigilance, and readiness to intervene. These processes are closely related to the latency of perceptual encoding and motor initiation, which are captured by the non-decision component of the DDM. Therefore, system reliability was linked to non-decision time rather than to drift rate or boundary setting. Drawing on previous literature, system reliability primarily modulates drivers' monitoring engagement and motor readiness, which directly influences the latency of detecting a takeover request and initiating manual control (Beller et al., 2013; Li et al., 2025; Schwarz et al., 2019). Accordingly, reliability is mapped onto the non-decision time, allowing this component to reflect system-level influences on drivers' readiness to act. The non-decision time is modeled as:

$$T_{\mathrm{er},i} = T_0 + b_{\mathrm{Rel}} \cdot \mathrm{Reliability}_i, \qquad (6)$$

where $T_0$ denotes the baseline non-decision time and $b_{\mathrm{Rel}}$ quantifies how perceived system reliability modulates perceptual–motor processing speed. This formulation captures systematic effects of perceived reliability on sensory–motor processing while maintaining interpretable and stable estimates within the drift–diffusion framework.

#### *4.2.4 Integrated Model Formulation*

The takeover decision process for driver is modeled as a drift–diffusion mechanism in which noisy evidence accumulates over time until reaching one decision boundary. The evolution of the decision variable $x_i(t)$ is described by:

$$\frac{dx_i(t)}{dt} = g_i + \varepsilon(t), \qquad (7)$$

Where $g_i$ is the drift rate defined in Equation (3), and $\varepsilon(t)$ denotes Gaussian white noise added to the accumulated evidence. A decision is triggered when the accumulated evidence reaches boundary $a_i(t)$, whose initial height and time-varying collapse are defined in Equations (4) and (5). The total response time consists of the decision time (i.e., the time at which the boundary is crossed) plus the non-decision time $T_{\mathrm{er},i}$ specified in Equations (6). Together, these equations fully specify the drift rate, collapsing boundary, initial decision threshold, and non-decision time that govern the takeover decision process in the modeled takeover scenario.

## 5. Result

### 5.1 Model Parameters Determining

We implemented the DDMs using the *pyddm* library and calibrated the model parameters to the empirical takeover data via a differential evolution (DE) optimization algorithm(Storn & Price, 1997). DE is a stochastic, population-based evolutionary algorithm designed for global optimization over continuous spaces. Unlike gradient-based methods, DE does not require the objective function to be differentiable and is highly robust against getting trapped in local minima, making it particularly suitable for the DDM. The fitting objective was the robust negative log-likelihood of the observed trial-level takeover response times. For each valid trial, the model was solved under the corresponding condition values, and the likelihood contribution was evaluated from the model-predicted first-passage-time distribution for the observed response time. The total fitting objective was obtained by summing the robust negative log-likelihood across all valid trials. Thus, the DDM fitting procedure used the full set of valid trial-level response-time observations rather than aggregated condition-level mean RTs. The resulting fitted parameters are summarized in Table 3.

**Table 3. The fitted DDM parameters.**

| **parameter** | $\alpha_0$ | $b_{speed}$ | $b_{flow}$ | $b_{day}$ | $B_0$ | $b_{experience}$ | k | $B_{min}$ | $t_0$ | $b_{reliance}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0.116 | 1.597 | 1.299 | -0.161 | 2.873 | 0.197 | 0.291 | 0.370 | 0.253 | 0.342 |

## 5.2 Model Comparison

The primary purpose of the model comparison was to evaluate the structural contribution of the proposed MDDM framework. Specifically, we compared a set of structurally ordered DDM-based models to examine whether hierarchical decomposition and factor-specific parameter mapping improved the modeling of takeover response times. This comparison served as a structural ablation analysis of the proposed framework. In addition, supplementary machine-learning benchmarks are reported in Appendix C. These models were included only as external predictive references and were not used as the primary basis for evaluating the theoretical contribution of the MDDM.

To evaluate the contribution of hierarchical decomposition and factor-specific parameter mapping to takeover decision modeling, eight model variants were specified and compared using mean absolute error (MAE) between simulated and empirical takeover response times. MAE was selected as the primary evaluation metric because it provides a direct and interpretable measure of predictive correspondence and is widely used in behavioral and driving-related modeling studies (Chen et al., 2023; Willmott & Matsuura, 2005). In addition, MAE avoids the disproportionate influence of extreme values commonly observed in takeover reaction time data. MAE was defined as the absolute difference between predicted and observed takeover times.

MAE was calculated at the trial level to preserve trial-by-trial variability in takeover response times and to avoid the loss of information that may result from participant-level or condition-level aggregation. For each trial, the fitted DDM generated a response-time distribution conditional on that trial's model inputs. The mean of this model-implied response-time distribution was used as the DDM-predicted takeover response time for that trial. The absolute error was then calculated as the absolute difference between the observed takeover response time and the DDM-predicted mean response time. The final MAE was obtained by averaging these absolute errors across all valid trials. No participant-level or condition-level aggregation was applied.

For clarity, the compared models were ordered by increasing structural complexity. Model 1 was a Standard DDM without conditioning on any experimental factors, serving as a minimal reference. Model 2 was a Direct-drift DDM (no-HLM), in which experimental factors were entered directly into the drift component without using HLM-derived weights or hierarchical decomposition. Model 3 was a Categorical DDM (no-HLM), in which experimental factors were directly mapped onto DDM parameters without hierarchical decomposition. Models 4–6 retained the hierarchical estimation stage but selectively incorporated only environmental (Model 4), human (Model 5), or system (Model 6) factors, respectively. As a structural control, Model 7 specified an MDDM variant in which all factors were mapped exclusively to the drift rate, allowing us to test whether performance gains require mapping factors to specific decision components. Model 8, the Full MDDM, integrated environmental, human, and system factors within a unified hierarchical framework by mapping each class of factors to the decision components they are theoretically expected to influence.

As shown in Table 4, model performance differed substantially across model structures. The Standard DDM (Model 1) exhibited relatively large MAE values, indicating a limited ability to capture systematic variation in takeover behavior when decision parameters were not conditioned on experimental factors. The Direct-drift DDM (no-HLM) (Model 2), in which experimental factors were entered directly into the drift component, slightly reduced MAE relative to the Standard DDM. This result suggests that forcing all experimental influences into the drift rate was insufficient. Introducing experimental factors directly into the DDM without hierarchical decomposition (Model 3) substantially reduced MAE relative to the standard DDM, indicating that incorporating experimental factors improves the model's ability to capture systematic variation in takeover response times. However, its performance remained inferior to the Full MDDM, suggesting that direct condition-to-parameter mapping alone is insufficient and that hierarchical decomposition combined with theoretically specified parameter mapping provides additional explanatory and predictive value.

Reductions in MAE were observed for the reduced MDDM variants (Models 4–6), indicating that retaining hierarchical estimation while selectively incorporating environmental, human, or system factors improves distributional correspondence. In contrast, mapping all factors exclusively to the drift rate (Model 7) did not lead to further reductions in MAE, indicating that performance gains depend on assigning experimental factors to appropriate decision components rather than concentrating them on a single parameter.

Among the DDM-based model variants, the Full MDDM (Model 8) achieved the lowest MAE, demonstrating that hierarchical estimation combined with factor-specific parameter mapping across distinct decision components provided the strongest model–data correspondence. Collectively, these results indicate that improvements in model performance depend not merely on incorporating more information into the decision process, but on correctly structuring how different classes of factors are mapped onto the decision components.

**Table 4. Comparison of diffusion model fits**

| Model | Model description | MAE(s) |
|---|---|---|
| Model 1 | Standard DDM (no factors) | 0.251 |
| Model 2 | Direct-drift DDM (no-HLM) | 0.235 |
| Model 3 | Categorical DDM (no-HLM) | 0.124 |
| Model 4 | MDDM (environmental factor only) | 0.161 |
| Model 5 | MDDM (human factor only) | 0.194 |
| Model 6 | MDDM (system factor only) | 0.222 |
| Model 7 | MDDM (all factors mapped to drift) | 0.216 |
| Model 8 | Full MDDM | 0.076 |

### 5.3 Simulation-Based Interpretation of the Fitted MDDM

For each condition, we performed Monte Carlo simulations of the evidence accumulation process (1,000 trajectories per condition) under the calibrated parameters. The purpose of these simulations was to examine how the fitted MDDM, with its theoretically specified parameter structure, generates observable takeover response-time distributions. In doing so, the simulations provide a process-level interpretation of how factor-related differences are expressed through different decision components, rather than an additional test of the pre-specified factor-to-parameter mappings. Figure 3 presents the resulting cumulative distributions of simulated takeover response times across the modeled conditions. The x-axis represents takeover response time (in seconds), and the y-axis represents cumulative probability. More specifically, changes in drift rate primarily affect the rate at which the cumulative probability rises. Faster evidence accumulation results in steeper curves that rise earlier in time. Changes in the decision boundary mainly produce a rightward shift of the cumulative distribution, reflecting that more evidence is required before initiating a takeover response. Changes in non-decision time primarily shift the onset of the distribution while leaving the overall shape of the curve largely unchanged.

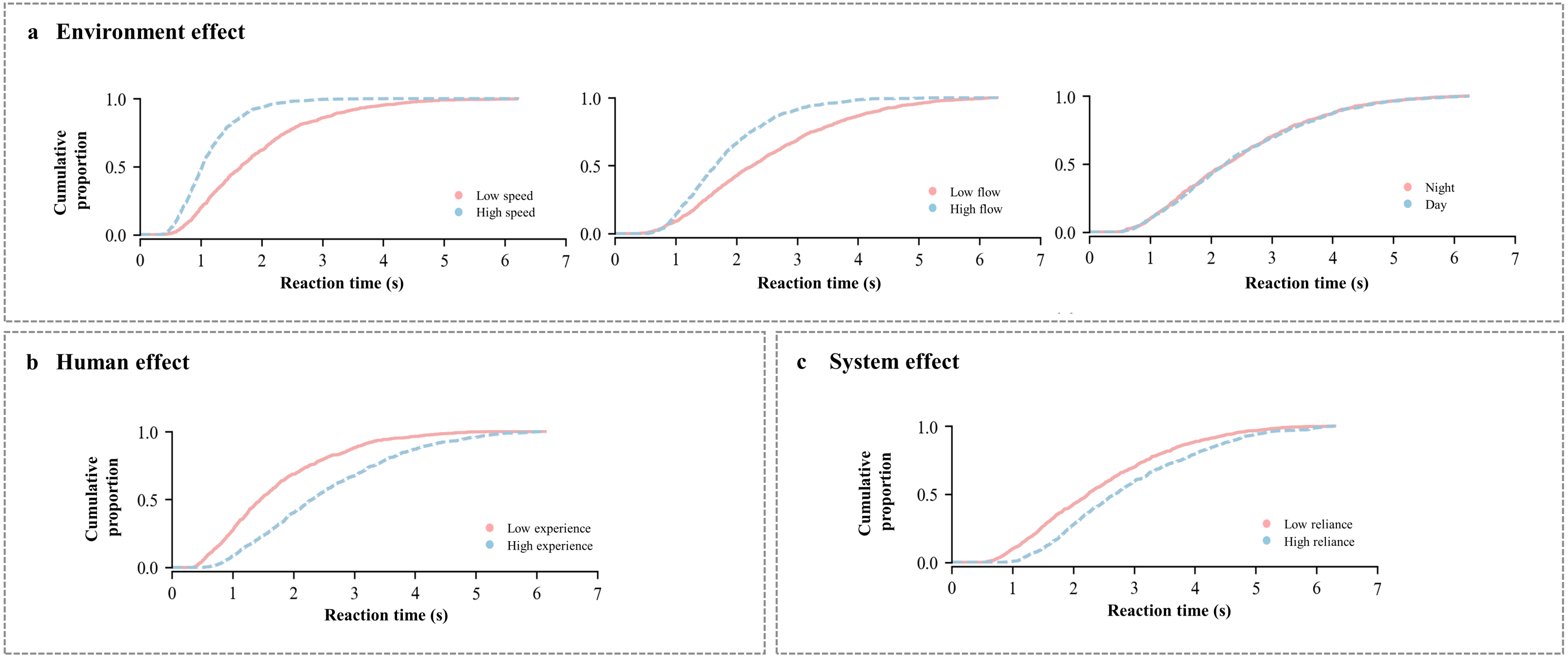


**Figure 3. DDM-simulated takeover decision process under different factors: (a) environmental factors (speed, optical flow, lighting), (b) human factor (driving experience), and (c) system factor (reliability).**

### *5.3.1 Environmental Factors and Evidence Accumulation*

As shown in Figure 3a, variations in environmental factors are captured in the fitted MDDM as changes in drift rate, which determines how rapidly risk-related evidence accumulates. Higher vehicle speed produces a steeper rise in the cumulative probability curve, reflecting a larger drift rate that drives faster movement of the decision variable toward the threshold. This pattern is consistent with the greater perceptual urgency induced by higher kinetic energy and shorter time-to-collision. Likewise, stronger optical flow cues also increase the drift rate, resulting in accelerated evidence accumulation and earlier takeover decisions. In contrast, daytime lighting conditions slightly reduce the drift rate, flattening the cumulative probability curve. Together, these results demonstrate that environmental factors modulate the rate at which evidence is accumulated, with higher-risk environments accelerating the decision process.

### *5.3.2 Human Factors and Decision Boundary*

As shown in Figure 3b, differences in driving experience are captured in the fitted MDDM as changes in the decision boundary, which determines how much evidence is required before a takeover response is initiated. The results indicate that evidence accumulates at a similar rate for both novice and experienced drivers. However, experienced drivers are modeled with a higher decision boundary than novices, meaning that more accumulated evidence is needed before a response is triggered. This produces a rightward shift in the cumulative probability curve, reflecting later takeover times despite an unchanged accumulation rate. Overall, these results demonstrate that driving experience modulates the decision threshold, with greater experience associated with a higher and more stable decision criterion for takeover.

#### *5.3.3 System Factors and Non-decision Time*

As shown in Figure 3c, differences in system reliability are captured in the fitted MDDM as changes in non-decision time, which reflects perceptual encoding and motor initiation processes outside evidence accumulation. The shape of the cumulative probability curve remains unchanged across different reliability conditions. Instead, reliability shifts the overall timing of responses: drivers in high reliability condition display longer non-decision times, which delays the onset of observable responses and produces a rightward shift in the early portion of the cumulative distribution. Conversely, lower system reliability shortens the non-decision interval, leading to an earlier rise in the cumulative curve. This pattern reflects the role of reliability in modulating drivers' monitoring engagement and readiness to act, influencing the latency of the response.

## 6. Discussion

This study provides a first empirical demonstration of the proposed MDDM framework in the context of takeover events when driving with automation. By integrating hierarchical modeling with evidence accumulation dynamics, we seek to bridge human, system, and environmental influences with cognitive parameters of decision-making. Specifically, we examine how environmental factors (e.g., speed, light), human factors (e.g., driving experience), and system factors (e.g., automation system reliability) contribute to distinct DDM parameters such as evidence accumulation, decision threshold setting, and non-decision processes. This approach offers a principled pathway toward modeling takeover decisions, advancing both psychological understanding and practical modeling of driving automation safety.

The proposed framework was also fitted using empirical takeover data, and the estimated parameters were able to reproduce the overall distributional characteristics of observed takeover responses. Simulated decision-time distributions showed good match with the empirical data at the aggregate level, indicating that the model captures core cognitive dynamics underlying takeover behavior. This alignment supports the plausibility of the framework and suggests that the estimated parameters reflect psychologically meaningful mechanisms rather than arbitrary numerical fits.

Beyond overall model fit, the model comparison results clarify why this framework is necessary. Performance improvements were found to depend on two complementary structural considerations. First, retaining a hierarchical estimation stage is essential for capturing systematic individual differences in takeover behavior; models that incorporated experimental factors without hierarchical decomposition showed limited gains. Second, beyond hierarchical structure alone, performance improvements further depended on how different classes of factors are mapped onto the decision process. Specifically, concentrating all influences on a single decision parameter was insufficient, whereas aligning environmental, human, and system factors with the decision components yielded the most accurate correspondence with empirical data. These findings suggest that effective modeling of takeover decisions requires both hierarchical representation of variability and structurally appropriate mapping of factors to distinct decision mechanisms.

Building on this structural foundation, the simulation-based analyses showed how environmental, human, and system-related factors are represented in different components of the takeover decision process. First, environmental cues are interpreted in the fitted model as differences in the speed of evidence accumulation. This finding aligns with accumulator models in manual driving (e.g., Durrani & Lee, 2023), which posit that kinematic cues directly determine the rate of evidence accrual. However, our results extend this mechanism to the context of driving automation, confirming that optical flow and lighting conditions serve as the critical perceptual evidence that informs urgency judgments, consistent with findings by Buzon et al. (2021) and Köhler et al. (2022; 2025).

Second, driving experience is interpreted in the fitted model as differences in boundary setting. This offers a nuanced perspective on the impact of experience reported in the literature. While prior studies have consistently demonstrated that experienced drivers exhibit better hazard anticipation and situation assessment (He et al., 2021; Sun & He, 2025), our modeling reveals that experienced drivers adopt more consistent and conservative criteria for initiating control actions, allowing them to avoid the erratic interventions often observed in novices.

Third, system reliability is interpreted in the fitted model as differences in non-decision processes, capturing changes in monitoring engagement and readiness to initiate action. This result provides a quantitative mechanism to explain the effects of reliability on driver behavior. As noted in previous research, high reliability can reduce monitoring engagement (Beller et al., 2013; Schwarz et al., 2019) and lower motor readiness (Li et al., 2025). Our framework confirmed that reliability primarily modulates the latency of detecting a request and initiating movement. Mapping these influences onto distinct DDM components provides a unified explanation for heterogeneous effects reported in previous takeover studies (Chen et al., 2021; Köhler et al., 2022; Schwarz et al., 2019; J. Wang et al., 2025; Zeeb et al., 2015).

From an applied perspective, the findings offer several implications for the design of takeover request strategies and driver support systems. First, as environmental factors govern the rate of evidence accumulation, takeover warnings should adapt to dynamic hazard information rather than rely solely on fixed distance or time thresholds. In other words, warning timing may benefit from being coupled to indicators of perceptual urgency, such as speed or optical flow, which directly influence evidence accumulation during takeover decisions. Second, the influence of driving experience on decision-boundary settings suggests that novice drivers may benefit not only from adaptive takeover support, but also from targeted training interventions aimed at stabilizing and better calibrating decision criteria. Because novice drivers exhibited lower decision boundaries, they may be more likely to initiate takeover actions under less stable decision strategies. Training protocols could therefore incorporate graded or escalating visual, auditory, or haptic warning cues to help novice drivers better calibrate urgency evaluation, maintain attentional readiness, and develop more stable takeover decision strategies under time pressure. Third, the effect of perceived system reliability on non-decision time highlights the importance of maintaining appropriate monitoring engagement while driving with automation to mitigate delayed response initiation.

## 7. Limitations and Future Work

Several limitations should be acknowledged. First, the present model assumes additive and independent contributions of environmental, human, and system-level factors, without incorporating potential nonlinearities or interactions. Although this simplification allows the cognitive parameters of the DDM to remain interpretable, it may overlook interaction effects. Future work may extend the framework to include interaction structures or nonlinear evidence accumulation dynamics. In addition, the dataset used to calibrate and validate the model was collected in a driving simulator, which offers experimental control but cannot fully reproduce the situational complexity and behavioral variability of real-world driving. Field studies or naturalistic driving data would provide an important next step for testing the generalizability and ecological robustness of the proposed MDDM framework.

## 8. Conclusion

This study proposed a multilevel drift–diffusion model (MDDM) framework for modeling takeover decisions in the context of driving automation. By decomposing the decision process and mapping environmental, human, and system factors onto distinct cognitive components, the model achieved good empirical fit while offering clear psychological interpretability. Specifically, environmental factors primarily shaped evidence accumulation, driving experience modulated decision boundaries, and system reliability influenced non-decision time. Beyond predictive performance, the proposed framework advances DDM-based modeling by integrating hierarchical structure with factor-specific parameterization in a complex human–automation interaction context. These findings suggest that cognitively grounded models such as the MDDM can support more explainable evaluation and design of driving automation systems and provide a transparent and extensible foundation for future research on adaptive control and driver state–aware interventions.

**Acknowledgements**

This research was supported by the National Nature Science Foundation of China (T2192932 and U2133209) and the Fundamental Research Funds of CAST (x242060302218).

**Reference**


Ayoub, J., Du, N., Yang, X. J., & Zhou, F. (2022). Predicting Driver Takeover Time in Conditionally Automated Driving. *IEEE Transactions on Intelligent Transportation Systems*, *23*(7), 9580-9589. https://doi.org/10.1109/TITS.2022.3154329

Banks, V. A., & Stanton, N. A. (2016). Keep the driver in control: Automating automobiles of the future. *Appl Ergon*, *53 Pt B*, 389-395. https://doi.org/10.1016/j.apergo.2015.06.020

Barisic, A., Sigrist, P., Oliver, S., Sciarra, A., & Winckler, M. (2023). *Driver Model for Take-Over-Request in Autonomous Vehicles* Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization, Limassol, Cyprus. https://doi.org/10.1145/3563359.3596994

Bathla, G., Bhadane, K., Singh, R. K., Kumar, R., Aluvalu, R., Krishnamurthi, R., Kumar, A., Thakur, R. N., & Basheer, S. (2022). Autonomous Vehicles and Intelligent Automation: Applications, Challenges, and Opportunities. *Mobile Information Systems*, *2022*(1), 7632892. https://doi.org/10.1155/2022/7632892

Beller, J., Heesen, M., & Vollrath, M. (2013). Improving the Driver–Automation Interaction:An Approach Using Automation Uncertainty. *Human Factors*, *55*(6), 1130-1141. https://doi.org/10.1177/0018720813482327

Buzon, L. G., Figueira, A. C., Larocca, A. P. C., & Oliveira, P. T. M. S. (2021). Effect of Speed on Driver's Visual Attention: A Study Using a Driving Simulator. *Transportation in Developing Economies*, *8*(1), 1. https://doi.org/10.1007/s40890-021-00139-y

Chen, F., Lu, G., Lin, Q., Zhai, J., & Tan, H. (2021). Are novice drivers competent to take over control from level 3 automated vehicles? A comparative study with experienced drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *81*, 65-81. https://doi.org/10.1016/j.trf.2021.05.012

Chen, H., Zhao, X., Li, H., Gong, J., & Fu, Q. (2024). Predicting driver's takeover time based on individual characteristics, external environment, and situation awareness. *Accident Analysis & Prevention*, *203*, 107601. https://doi.org/https://doi.org/10.1016/j.aap.2024.107601

Chen, L., Wu, P., Chitta, K., Jaeger, B., Geiger, A., & Li, H. (2024). End-to-End Autonomous Driving: Challenges and Frontiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *46*(12), 10164-10183. https://doi.org/10.1109/TPAMI.2024.3435937

Chen, X., Zhu, M., Chen, K., Wang, P., Lu, H., Zhong, H., Han, X., Wang, X., & Wang, Y. (2023). FollowNet: A Comprehensive Benchmark for Car-Following Behavior Modeling. *Scientific Data*, *10*(1), 828. https://doi.org/10.1038/s41597-023-02718-7

Du, N., Zhou, F., Pulver, E. M., Tilbury, D. M., Robert, L. P., Pradhan, A. K., & Yang, X. J. (2020). Predicting driver takeover performance in conditionally automated driving. *Accident Analysis & Prevention*, *148*, 105748. https://doi.org/https://doi.org/10.1016/j.aap.2020.105748

Durrani, U., & Lee, C. (2023). Applying the Accumulator model to predict driver's reaction time based on looming in approaching and braking conditions. *Journal of Safety Research*, *86*, 298-310. https://doi.org/https://doi.org/10.1016/j.jsr.2023.07.008

Gipps, P. G. (1981). A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*, *15*(2), 105-111. https://doi.org/10.1016/0191-2615(81)90037-0

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1938-1942. https://doi.org/10.1177/1541931213571433

Hasegawa, K., Wu, Y., & Kihara, K. (2024). Two-stage transition procedure reduces potential hazards on planned transitions in automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *107*, 924-936. https://doi.org/10.1016/j.trf.2024.10.017

He, D., DeGuzman, C. A., & Donmez, B. (2021). Anticipatory Driving in Automated Vehicles: The Effects of Driving Experience and Distraction. *Human Factors*, *65*(4), 663-663. https://doi.org/10.1177/00187208211026133

Holley, D., sa, J. D., Mahjoub, H. N., Ali, G., Chalaki, B., & Moradi-Pari, E. (2023, 24-28 Sept. 2023). MR-IDM - Merge Reactive Intelligent Driver Model: Towards Enhancing Laterally Aware Car-following Models. 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC),

Horswill, M. S. (2016). Hazard Perception in Driving. *Current Directions in Psychological Science*, *25*(6), 425-430. https://doi.org/10.1177/0963721416663186

Hu, J., Sun, S., & Wang, R. (2022). Research on the influence of light source characteristics on traffic visual distance in foggy areas at night. *Building and Environment*, *212*, 108818. https://doi.org/10.1016/j.buildenv.2022.108818

Hu, T., Leppänen, I., & Franco, L. A. (2025). Intuition in Decision Making: Insights From Drift Diffusion Modeling. *Journal of Behavioral Decision Making*, *38*(3), e70033. https://doi.org/10.1002/bdm.70033

Huang, H., Li, Z., Cheng, H., Wang, H., Jiang, J., Li, X., & Zgonnikov, A. (2025). Understanding driver cognition and decision-making behaviors in high-risk scenarios: A drift diffusion perspective. *Accident Analysis & Prevention*, *220*, 108123. https://doi.org/https://doi.org/10.1016/j.aap.2025.108123

Ju, U., & Wallraven, C. (2023). Dynamic measurements of speed and risk perception during driving: Evidence of speed misestimation from continuous ratings and video analysis. *PLOS ONE*, *18*(9), e0291043. https://doi.org/10.1371/journal.pone.0291043

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.). *Decision making in action: Models and methods*, (pp. 138–147). Ablex.

Köhler, A.-L., Klatt, M., Koch, I., & Ladwig, S. (2023). Investigating the influence of visuospatial stimuli on driver's speed perception: a laboratory study. *Cognitive Research: Principles and Implications*, *8*(1), 59. https://doi.org/10.1186/s41235-023-00513-x

Köhler, A.-L., Koch, I., & Ladwig, S. (2022). Guiding drivers towards safer driving speed: Exploiting visual dominance in speed adaptation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *90*, 438-450. https://doi.org/10.1016/j.trf.2022.09.011

Köhler, A.-L., Koch, I., & Ladwig, S. (2025). Investigating the role of visual and corresponding auditory stimuli in driving-related speed perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, *110*, 1-14. https://doi.org/10.1016/j.trf.2025.02.007

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

Lee, M., & Pitts, B. J. (2024). The Impact of System Confidence and Reliability on Drivers' Decision-Making in Conditionally Automated Vehicles. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *68*(1), 915-920. https://doi.org/10.1177/10711813241260388

Li, J., Ye, Y., Liu, L., & Butz, A. (2025). Effect of Reliability, Its Framing and Error Bias on Trust in Human-Vehicle Collaboration. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *35*(3), e70008. https://doi.org/10.1002/hfm.70008

Liu, Q., Liu, J., Cai, Y., & Chen, L. (2022). Exploring the Impact of the Takeover Time for Conditionally Automated Driving Vehicles on Traffic Flow in Highway Merging Area. *IEEE Transactions on Intelligent Transportation Systems*, *23*(12), 24753-24764. https://doi.org/10.1109/TITS.2022.3202834

Miller, J. A., Nikan, S., & Zaki, M. H. (2024). Navigating the Handover: Reviewing Takeover Requests in Level 3 Autonomous Vehicles. *IEEE Open Journal of Vehicular Technology*, *5*, 1073-1087. https://doi.org/10.1109/OJVT.2024.3443630

Miqdady, T., Oña, R. d., Casas, J., & Oña, J. d. (2023). Studying Traffic Safety During the Transition Period Between Manual Driving and Autonomous Driving: A Simulation-Based Approach. *IEEE Transactions on Intelligent Transportation Systems*, *24*(6), 6690-6710. https://doi.org/10.1109/TITS.2023.3241970

Mohammad, S. H. A., Farah, H., & Zgonnikov, A. (2023, 24-28 Sept. 2023). Modeling Gap Acceptance in Overtaking: A Cognitive Process Approach. 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC),

Morales-Alvarez, W., Sipele, O., Léberon, R., Tadjine, H. H., & Olaverri-Monreal, C. (2020). Automated Driving: A Literature Review of the Take over Request in Conditional Automation. *Electronics*, *9*(12), 2087. https://www.mdpi.com/2079-9292/9/12/2087

Myers, C. E., Interian, A., & Moustafa, A. A. (2022). A practical introduction to using the drift diffusion model of decision-making in cognitive psychology, neuroscience, and health sciences. *Front Psychol*, *13*, 1039172. https://doi.org/10.3389/fpsyg.2022.1039172

Pakdamanian, E., Sheng, S., Baee, S., Heo, S., Kraus, S., & Feng, L. (2021). *DeepTake: Prediction of Driver Takeover Behavior using Multimodal Data* Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan. https://doi.org/10.1145/3411764.3445563

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, *20*(4), 260-281.

https://doi.org/10.1016/j.tics.2016.01.007
Schwarz, C., Gaspar, J., & Brown, T. (2019). The effect of reliability on drivers' trust and behavior in conditional automation. *Cognition, Technology & Work*, *21*(1), 41-54. https://doi.org/10.1007/s10111-018-0522-y
Sheng, Y., Li, X., Sun, X., He, D., & Zhang, J. (2025). Measuring perceived risk in intelligent system interaction: Development and validation of a multidimensional scale across three safety-critical domains. https://doi.org/10.31234/osf.io/f7e3p_v1
Sheng, Y., Li, X., Sun, X., & Jingyu, Z. (2025). Measuring perceived risk in intelligent system interaction: Development and validation of a multidimensional scale across three safety-critical domains. https://doi.org/10.31234/osf.io/f7e3p_v1
Sheykhfard, A., Qin, X., Shaaban, K., & Koppel, S. (2022). An exploration of the role of driving experience on self-reported and real-world aberrant driving behaviors. *Accident Analysis & Prevention*, *178*, 106873. https://doi.org/10.1016/j.aap.2022.106873
Soares, S., Lobo, A., Ferreira, S., Cunha, L., & Couto, A. (2021). Takeover performance evaluation using driving simulation: a systematic review and meta-analysis. *European Transport Research Review*, *13*(1), 47. https://doi.org/10.1186/s12544-021-00505-2
Storn, R., & Price, K. (1997). Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces. *Journal of Global Optimization*, *11*(4), 341-359. https://doi.org/10.1023/A:1008202821328
Sun, M., & He, D. (2025). The influence of cognitive distractions and driving experience on hazard perception performance during partially automated driving. *Accident Analysis & Prevention*, *221*, 108202. https://doi.org/10.1016/j.aap.2025.108202
Tan, X., & Zhang, Y. (2022). The effects of takeover request lead time on drivers' situation awareness for manually exiting from freeways: A web-based study on level 3 automated vehicles. *Accident Analysis & Prevention*, *168*, 106593. https://doi.org/10.1016/j.aap.2022.106593
Treiber, M., Hennecke, A., & Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, *62*(2), 1805-1824. https://doi.org/10.1103/PhysRevE.62.1805
Wang, C., Ren, W., Xu, C., Zheng, N., Peng, C., & Tong, H. (2025). Exploring the Impact of Conditionally Automated Driving Vehicles Transferring Control to Human Drivers on the Stability of Heterogeneous Traffic Flow. *IEEE Transactions on Intelligent Vehicles*, *10*(2), 912-928. https://doi.org/10.1109/TIV.2024.3419789
Wang, J., Li, J., Xiao, Y., & Song, K. (2025). Steering in the dark: The impact of environmental luminance on driver behavior through optical flow analysis. *Attention, Perception, & Psychophysics*, *87*(8), 2417-2429. https://doi.org/10.3758/s13414-025-03146-y
Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*, 79-82. https://www.int-res.com/abstracts/cr/v30/cr030079
Wood, J. M. (2020). Nighttime driving: visual, lighting and visibility challenges. *Ophthalmic and Physiological Optics*, *40*(2), 187-201. https://doi.org/10.1111/opo.12659
Wu, H., Wu, C., Lyu, N., & Li, J. (2022). Does a faster takeover necessarily mean it is better?

A study on the influence of urgency and takeover-request lead time on takeover performance and safety. *Accident Analysis & Prevention*, *171*, 106647. https://doi.org/10.1016/j.aap.2022.106647

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. *Accid Anal Prev*, *78*, 212-221. https://doi.org/10.1016/j.aap.2015.02.023

Zhou, B., Feng, Z., Liu, J., Huang, Z., & Gao, Y. (2024). A method to enhance drivers' hazard perception at night based on "knowledge-attitude-practice" theory. *Accident Analysis & Prevention*, *200*, 107565. https://doi.org/10.1016/j.aap.2024.107565

**Appendix A**

**Table A1. Search ranges and constraints for key DDM parameters during differential evolution optimization**

| Parameter | Description | Search range |
|---|---|---|
| $\alpha_0$ | Baseline drift term | [-2, 2] |
| $b_{speed}$ | Speed-related coefficient | [-2, 2] |
| $b_{flow}$ | Optical-flow-related coefficient | [-2, 2] |
| $b_{day}$ | Lighting-related coefficient | [-2, 2] |
| $B_0$ | Initial boundary height | [0.6, 3] |
| $b_{experience}$ | Experience-related boundary coefficient | [-0.6, 0.6] |
| k | Boundary collapse-rate parameter | [0.0, 1.0] |
| $B_{min}$ | Minimum boundary level | [0.1, 0.6] |
| $t_0$ | Baseline non-decision time | [0.02, 0.60] |
| $b_{reliance}$ | Reliability-related coefficient on non-decision time | [-0.6, 0.6] |

**Table A2. Bootstrap uncertainty intervals for the MDDM parameters**

| Parameter | 95% bootstrap uncertainty interval |
|---|---|
| $\alpha_0$ | [0.041, 0.278] |
| $b_{speed}$ | [1.116, 1.923] |
| $b_{flow}$ | [0.892, 1.667] |
| $b_{day}$ | [-0.402, -0.083] |
| $B_0$ | [2.451, 2.982] |
| $b_{experience}$ | [0.018, 0.374] |
| k | [0.102, 0.486] |
| $B_{min}$ | [0.246, 0.488] |
| $t_0$ | [0.166, 0.349] |
| $b_{reliance}$ | [0.104, 0.542] |

**Appendix B. The final Version of Multidimensional Perceived Risk Scale (MPRS)**

The final version of the scale consists of 27 items rated on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). Scores for each dimension can be computed by averaging the items belonging to the corresponding dimension. An overall perceived risk score can also be obtained by averaging all items, with higher scores indicating higher levels of perceived risk. Depending on the research objectives and interaction context, researchers may use selected dimensions of the MPRS independently, rather than administering the full set of dimensions. The specific items and dimensional structure of the MPRS are reported below.

Third-Party Risk (TPR) This dimension captures users' concerns that interacting with the intelligent system may negatively affect other people.

1. Using the system may expose others to potential risk.
2. Using the system may cause trouble for others.
3. Using the system may cause others to suffer losses.
4. Using the system may impose unnecessary consequences on others.
5. Using the system may endanger the safety of others.

Operational Cost Risk (OCR) This dimension reflects users' perceptions that interacting with the system leads to inefficiencies, such as wasted time or effort.

6. Using the system wastes my time.
7. Using the system reduces my task efficiency.
8. Using the system increases my time pressure.
9. Using the system slows down my task completion.

10. Using the system wastes my energy.

Social Appraisal Risk (SAR) This dimension captures concerns about being negatively evaluated by others because of using the intelligent system.

11. Using the system may affect how others evaluate my abilities.

12. Using the system makes me worry about how others perceive me.

13. Using the system makes me worry that others may question my abilities.

14. Using the system makes others think I am not cautious or responsible.

15. Using the system makes others think I am foolish.

Probability and Severity Risk (PSR) This dimension reflects users' cognitive appraisal of the likelihood and potential severity of negative outcomes resulting from system failure.

16. When using the system, I worry that the system's stability cannot be guaranteed.

17. When using the system, I feel the probability of system errors is high.

18. When using the system, I think the risk of system errors is significant.

19. Using this system is generally risky.

20. Using the system may cause me some kind of loss.

Affective Risk (AR) This dimension represents emotional discomfort, such as anxiety or nervousness, experienced during interaction with the system.

21. Using this system makes me feel anxious.

22. Using this system makes me feel nervous.

23. Using this system makes me feel scared.

24. Using this system makes me feel psychologically uncomfortable.

Performance Integrity Risk (PIR) This dimension assesses users' doubts about whether the intelligent system can reliably and effectively accomplish intended tasks.

25. The system is reliable in this scenario.

26. This system allows me to complete the tasks I expect.

27. The system can accurately meet my needs in the current task.

Note: Items in the Performance Integrity Risk (PIR) dimension were reverse coded such that higher scores indicate higher perceived risk.

**Appendix C. Supplementary machine-learning benchmark**

To provide an external predictive reference, we additionally evaluated three commonly used machine-learning regression models: Random Forest, XGBoost, and LightGBM. These models were included as supplementary benchmarks only and were not used as the primary basis for evaluating the theoretical contribution of the proposed MDDM. The main model-comparison analysis in the manuscript focuses on the structural ablation among DDM-based model variants.

The machine-learning models used the same experimental condition variables as input features: vehicle speed, optical flow, lighting condition, driving experience, and system reliability. For the machine-learning baselines, MAE was calculated from trial-level cross-validated point predictions. No condition-level aggregation was applied to the machine-learning predictions. Thus, both model families used the same trial-level response variable as the prediction target.

**Table C1. Hyperparameter configurations of the machine learning baseline models.**

| Model | Parameter / Specification | Value / Description |
|---|---|---|
| Common Settings | Validation Strategy | 5-Fold Within-Subject Cross-Validation |
| | Input Features (N=5) | Speed, Flow, Lighting, Experience, System Reliability |
| | Software Environment | Python 3.9; scikit-learn (v1.0.2); NumPy (v1.21.2) |
| | Random Seed | Fixed to 42 for reproducibility |
| Random Forest | Number of Trees | 400 |
| | Maximum Tree Depth | None (Unrestricted) |
| | Min. Samples per Leaf | 2 |
| | Max. Features | Square root of total features (p) |
| | Implementation Pkg. | scikit-learn (v1.0.2) |
| XGBoost | Number of Trees | 400 |
| | Learning Rate | 0.05 |
| | Maximum Tree Depth | 3 |
| | Objective Function | Squared Error (Regression) |

| | | |
|---|---|---|
| | Implementation Pkg. | xgboost (v1.5.0) |
| LightGBM | Number of Trees | 400 |
| | Learning Rate | 0.05 |
| | Maximum Tree Depth | 3 |
| | Boosting Type | GBDT (Gradient Boosting Decision Tree) |
| | Implementation Pkg. | lightgbm (v3.3.2) |

The machine-learning baselines produced higher MAE values under the present evaluation setting. However, these results should be interpreted cautiously because the benchmark models were based on direct point prediction from a limited set of experimental condition variables. In contrast, the DDM-based models were designed to model response-time distributions and to provide cognitively interpretable decision-process parameters. Therefore, the comparison is intended to provide both a structural evaluation of the proposed MDDM and an external predictive benchmark, rather than a purely algorithmic competition.



**Table C2. Comparison of diffusion model fits**

| **Model** | **Model description** | **MAE(s)** |
|---|---|---|
| Model 9 | RandomForest | 0.7050 |
| Model 10 | XGBoost | 0.7084 |
| Model 11 | LightGBM | 0.7082 |