Knowing When to Move:

Generalisable Kinematics-Dependent Drift Diffusion Modelling of Human Narrow Passage Interactions

ME51035: MSc Thesis

F.R.A. van Waveren

Delft University of Technology

TU Delft

Knowing When to Move:

Generalisable Kinematics-Dependent Drift Diffusion Modelling of Human Narrow Passage Interactions

by

F.R.A. van Waveren

Defense date

15 January 2026

Responsible supervisor:	Dr. A. Zgonnikov
Daily Supervisor:	Dr. B. Nallapu
Examiner:	Dr. D. Dodou
Project Duration:	April 2025 - January 2026
Faculty:	Faculty of Mechanical Engineering

Cover: Cover image generated by the author using an AI-based image generation tool (prompted and curated by the author)

TU Delft

Deciding When to Move: Generalisable Kinematics-Dependent Drift Diffusion Modelling of Human Narrow Passage Interactions

Author: F.R.A. van Waveren
(https://orcid.org/0009-0002-5385-0335)
Responsible Supervisor: Dr. A. Zgonnikov
Daily Supervisor: Dr. B.T. Nallapu

Faculty of Mechanical Engineering
Department of BioMechanical Engineering & Department of Cognitive Robotics,
Delft University of Technology
Delft, Netherlands

Defended on: 15 January 2026

Abstract— Human drivers routinely make interaction decisions by integrating multimodal, noisy perceptual information over time to decide whether to proceed or yield. Understanding the cognitive mechanisms underlying such decisions is crucial for explaining driver behaviour and for developing models that generalise across traffic scenarios. This study examines human driver decision-making in a narrow-passage gap-acceptance task, an interaction scenario characterised by ambiguity of priority and described by interdependent kinematic and visual information. Behavioural data from laboratory ($N = 36$) and ($N = 175$) online experiments were analysed to assess whether decision dynamics generalise across populations with differing variability and experimental control. Across datasets, decision behaviour showed systematic dependencies on kinematic conditions, with increased reaction times (RTs) and choice variability in ambiguous situations. Longitudinal kinematics-based drift diffusion models (LK-DDMs) captured both decision proportions and reaction-time distributions in the narrow passage task and generalised across datasets with differing coverage of the experimental conditions, including lab and online data. The same accumulation framework transferred to a related decision-making task, indicating that the inferred dynamics are not task-specific. Incorporating visual looming in the drift function yielded selective improvements in short-distance *Wait* decisions, without global gains in reaction-time accuracy.

---

1 Introduction

1.1 Trends in driver decision modelling

Human drivers continuously act under uncertainty, integrating sensory information and motor control in a dynamic and time-pressured environment. Understanding and predicting such behaviour is essential for improving traffic safety, enabling automated vehicles to interact smoothly with human drivers, and advancing theories of human sensori-motor decision-making. However, many applied models of driver behaviour are rooted in machine-learning or rule-based [1], [2], [3], [4], [5], focusing on reproducing observed behaviour rather than explaining the cognitive processes underlying decision variability, hesitation, and context-dependent timing [6], [7]. Evidence-accumulation models (EAMs) provide a mechanistic alternative by linking observable choices and RTs to latent decision dynamics. By modelling decisions as the gradual accumulation of noisy evidence toward a response threshold, Drift Diffusion Models (DDMs) offer a unified account of decision outcomes and response timing [8]. Such models have been increasingly applied to driving tasks, including braking, steering control, gap acceptance, and interaction decisions under uncertainty [9], [10], [11], [12], [13]. Despite this growing body of work, applications remain largely task-specific, and it remains unclear which modelling principles generalise across interaction scenarios, datasets, and populations.

1.2 The narrow passage task as an interaction decision problem

Narrow-passage interaction decisions–scenarios in which two road users approach a narrow road section that cannot be traversed by both vehicles at the same time–provide a particularly informative setting for studying driver decision-making. In many such scenarios, drivers have equal right of way, resulting in structurally ambiguous situations in which neither proceeding nor yielding is obviously correct. While narrow passages have often been studied from cooperative or communication-focused perspectives [14], [15], the baseline decision dynamics of individual drivers in the absence of explicit coordination remain less well understood. Decisions in the narrow-passage task require drivers to integrate multiple streams of information, including visual angles, distance, velocity, and time-to-arrival, under time pressure and interactional ambiguity. This makes the task well-suited for examining how kinematic cues shape both choice and response timing in longitudinal interactions. Although evidence-accumulation models have been applied to related interaction decisions, such as gap acceptance and pedestrian crossing [11], [16], no systematic EAM-based modelling of narrow-passage decisions currently exists, and no prior work has examined whether inferred decision dynamics generalise across datasets or task variants.

1.3 Drift diffusion models for driver decision-making

DDMs originate in mathematical psychology as a mechanistic framework for two-alternative forced-choice decisions under uncertainty [8]. In a DDM, noisy evidence is accumulated over time until it reaches one of two decision boundaries, corresponding to alternative actions. The rate of accumulation (the drift) captures how strongly the available evidence favours one decision outcome over the other, while variability in both choices and RTs is modelled as Gaussian noise. In driving contexts, a wide range of EAMs have been adapted to represent how perceptual information shapes interaction decisions. In driver decision modelling, drift terms are commonly defined as either functions of kinematic variables, including distance, velocity, or time-to-arrival, or as functions of visual-perceptual cues such as visual angle, optical expansion, and looming rate. Kinematics-dependent drift formulations have been used to model gap acceptance, pedestrian crossing, overtaking, and other interaction-level decisions [16], [17], [18], [19], while visual-perceptual formulations are more often applied to time-critical behaviours such as braking and collision avoidance [13], [20], [21]. Together, this literature highlights both the promise of DDMs for modelling driver decisions and the open question of how kinematic and visual information should be represented within a unified accumulation framework.

1.4 Aim and scope of the present study

The present study addresses three gaps in the existing literature. First, it provides the first systematic DDM-based analysis of narrow-passage approach decisions, a common yet under-modelled interaction scenario characterised by interactional ambiguity and competing kinematic cues. Second, it directly compares alternative kinematics-dependent drift formulations based on time-to-arrival, distance, and velocity, and examines the extent to which visual motion cues, such as looming, refine or alter the inferred decision dynamics. Third, it evaluates the robustness and generalisability of the learned accumulation dynamics through partial training data exclusion and cross-dataset validation, as well as the generalisability of the model framework across different tasks by refitting.

2 Methodology

This section describes the experimental setup (Section 2.1), the derivation of kinematic and visual predictors from video stimuli (Section 2.2), and the statistical analysis of decision patterns (Section 2.3). It then specifies the computational modelling framework and fitting procedure (Sections 2.4.2–2.4.9), followed by the validation strategies used to assess model generalisation and robustness (Section 2.5).

2.1 Experimental setup

2.1.1 Narrow passage gap acceptance scenario

The primary dataset analysed in this study was taken from a paper in preparation by Miller et al. [15]. During the experiment, participants viewed short (3 s) ego-perspective driving videos depicting an approach to a narrow passage where only one vehicle could proceed at a time. The narrow passage task is characterised by interactional ambiguity caused by the equal lawful right of way between the two agents [14]. To resolve this issue, participants

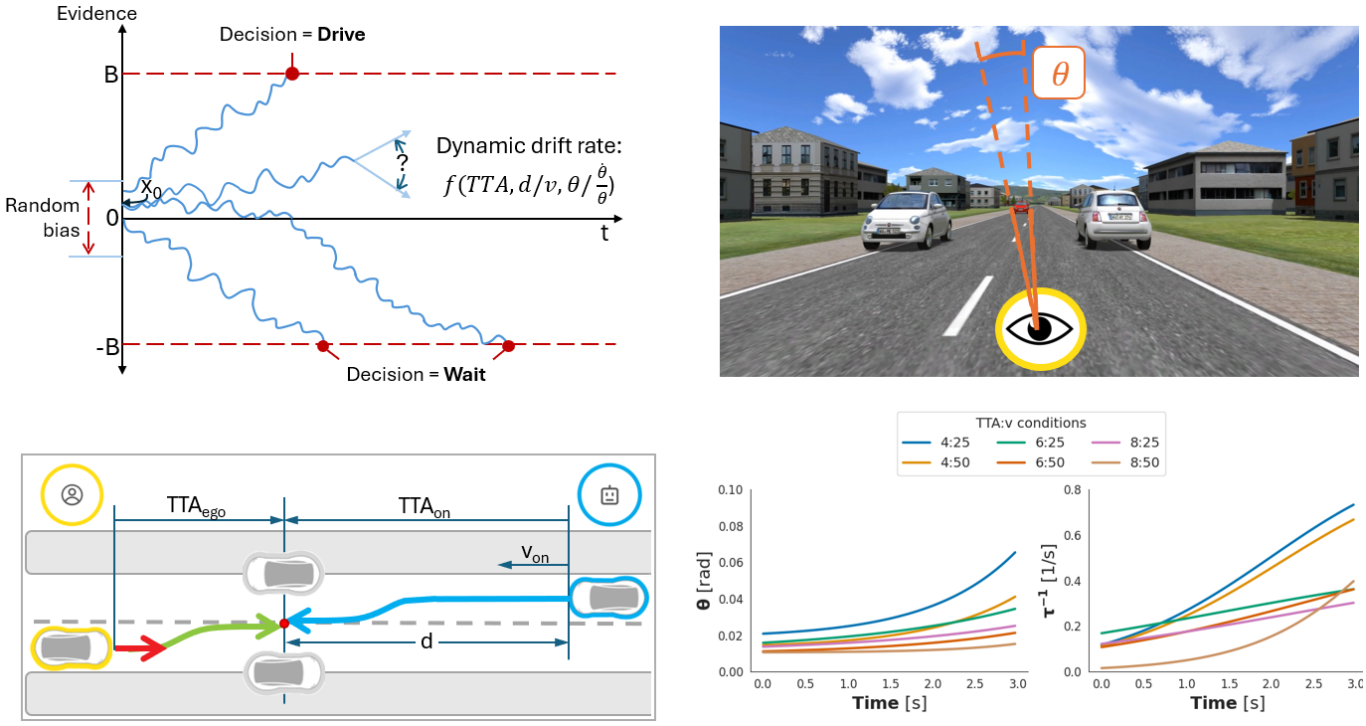


Figure 1: **Top left:** Schematic illustration of the decision variable dynamics in the drift–diffusion modelling framework used in this study. Evidence accumulates toward *Drive* (upper boundary) or *Wait* (lower boundary), with drift rate determined by kinematic variables and optional visual variables. **Top right:** Video frame from the narrow passage experiment, showing the oncoming vehicle (red) and indicating schematically the visual angle (θ) derived from the video. **Bottom left:** Schematic representation of the narrow passage scenario geometry, including ego and oncoming vehicle trajectories, distance d, initial velocity v_{on}, and time-to-arrival (TTA) variables. **Bottom right:** Visual angle and visual looming over time for experimental $\mathrm{TTA}_0 \times v_{on}$ conditions.

must choose a strategy based on only the information available from the videos. In each trial, participants indicated as quickly as possible whether they would *Drive* through the passage or *Wait* for the oncoming vehicle, responding via a keyboard key press.

The visual and geometric structure of the task is illustrated in Fig. 1. The top-right panel shows a representative ego-perspective video frame as presented to participants, while the bottom-left panel provides a top-down schematic of the same interaction, highlighting the spatial configuration of the ego vehicle, the oncoming vehicle, and the constrained road segment.

The videos were rendered in SILAB 6.5 at 30 FPS and presented without sound. Across all trials, participants observed an oncoming vehicle approaching the narrow passage from the opposite direction. The oncoming vehicle was either manually driven or automated; automated vehicles were visually marked using either a turquoise eHMI band or a red vehicle colour. The underlying kinematic scenario (i.e., positions and motions of all objects) and optical perspective were identical across trials. The automation and eHMI categories were not considered in the results and analyses of this paper, as they have almost no effect on response measures, according to Miller et al. [15].

2.1.2 Lab and online datasets

The narrow passage experiment was conducted in two independent experimental contexts. In the *lab dataset*, participants completed the task in a supervised environment using consistent display hardware, with screen distance partly controlled but dependent on individual physical characteristics. In the *online dataset*, participants completed the same task remotely using their own hardware in an unsupervised environment, resulting in uncontrolled display characteristics but a substantially larger sample size.

This dual-collection design enables analysis of behavioural responses under both controlled lab conditions and large-sample online conditions. Importantly, it allows the robustness and generalisability of subsequent modelling results to be assessed across populations and viewing environments that differ in experimental control.

Table 1 details the condition grid and corresponding starting distances of the oncoming vehicle, denoted by d in Figure 1, and Table 2 summarises the key characteristics of the lab and online datasets.

Table 1: Condition grid showing oncoming vehicle starting distances as a function of oncoming vehicle speed v_on and its initial time-to-arrival (TTA_0).

v_{on} \ TTA_0	4s	6s	8s
25 km/h	27.8m	41.7m	55.6m
50 km/h	55.6m	83.3m	111.1m

2.2 Derivation of visual angles from video

To ensure consistent preprocessing across both datasets, all trials were transformed into a common set of kinematic

Table 2: Overview of dataset demographics, experimental apparatus, design, independent and dependent variables.

	Lab dataset	**Online dataset**
Setting	Laboratory	Online
N participants	36	175
N trials	36×240	175×240
Age (μ/SD/range)	25.9/9.7/18–66	53.9/13.4/20–79
Gender (M/F)	14 / 22	118 / 57
Apparatus	30FPS SILAB-based simulated video 27" display Keyboard input	30FPS SILAB-based simulated video Unknown display Keyboard input
Design	3×2×2 within-subject	3×2×4 within-subject
Independent variables	TTA_0 [s]: (4, 6, 8) v_{on} [km/h]: (25, 50) Automation: (Y, N)	TTA_0 [s]: (4, 6, 8) v_{on} [km/h]: (25, 50) Automation: (Y, N, Y+eHMI, N+eHMI)
Dependent variables	Response time Decision outcome Demographic data Expectation certainty (post-questionnaire)	Response time Decision outcome Demographic data Expectation certainty (post-questionnaire)

and visual predictors suitable for subsequent statistical and computational modelling. For the narrow passage data, this required reconstructing perceptual variables from the raw stimuli, as the experimental input consisted of 30 FPS ego-perspective videos rather than direct simulator state variables.

A computer-vision-based approach was adopted to extract visual information directly from the rendered video frames. This choice avoids reliance on highly specific simulator metadata (e.g., exact object geometry or camera pose), which was not uniformly available, and makes the preprocessing pipeline easier to apply or adapt to new experiments and scenarios.

Visual angles were computed to reflect the optical information available to participants, rather than properties of the video stimulus itself. This approach has been widely used in models of driver behaviour in emergency manoeuvring [13], [20], [21]. Specifically, all visual angles were reconstructed from the estimated position of the participant's eye, assumed to be centred between the eyes, using known experimental setup dimensions and assumptions based on average German anatomical measurements, [22]. The required geometric relationships were reconstructed using calibration information and construction lines inferred from the only available photograph of the experimental setup. Detailed assumptions, derivations, and validation procedures are provided in Appendix A.

The transformation pipeline consisted of three main steps. First, the oncoming vehicle was detected in each frame using a deep-learning-based object detection algorithm (YOLOv8, [23]), yielding a time series of bounding-box dimensions. Second, the bounding-box width was converted into an instantaneous visual angle $\theta(t)$ using the calibrated projection geometry. Third, visual looming rate was computed as $\tau^{-1}(t) = \frac{\dot{\theta}(t)}{\theta(t)}$.

The bottom-right panel of the visual abstract (Fig. 1) shows how visual angle and visual looming evolve over time, as extracted from the narrow passage stimuli, highlighting their characteristic non-linear growth. These properties motivate their consideration as perceptual predictors in the statistical and computational analyses of this study.

Due to the constant speeds of both vehicles, the time series of the visual angle and looming must be smooth and strictly increasing. Therefore, exponential functions were fitted to the time series of bounding-box dimensions. For the implementation of visual information in the model-fitting pipeline, logistic functions were fitted to the resulting visual angle and looming trajectories. Alternative functions were evaluated at both stages but were rejected due to inferior fit.

2.3 Statistical analysis of decision patterns

Before applying cognitive modelling, the experimental data were evaluated qualitatively based on descriptive statistics and visualisation. Additionally, statistical regressions were performed to describe the effect sizes and reliability of the observed behavioural patterns. Specifically, the statistical analysis was conducted to determine the population-level contributions of kinematic and perceptual predictors to both decision outcome and RT, while accounting for participant-specific strategies or baseline differences.

2.3.1 Regression analysis design

Regression analysis was designed to identify population-level structure and account for participant-level variability present in RT and decision bias in the narrow passage datasets. To this end, a *mixed-effects* regression framework was adopted. Fixed effects capture systematic influences of task variables on the decision outcome and RT at the population level, while random effects account for stable participant-level differences.

Including random intercepts for participant ID significantly improved explanatory power for both decision outcome and RT analyses across both datasets. Exploratory analysis revealed substantial between-participant variability in baseline behaviour by determining the proportion of total variance in the data attributable to stable between-participant differences. Across regressions for both datasets, the attributed proportion of variance ranged from 0.18 to 0.37 (see Appendix C). This confirms that individual differences account for a considerable portion of the observed variance and justifies the use of participant-level random effects.

Models that additionally included random slopes for predictors such as TTA_0, τ^{-1}, or v_{on} resulted in singular or non-convergent fits, indicating insufficient evidence for systematic between-participant differences in these effects.

Consequently, only participant-level random intercepts were retained in the final mixed-effects models.

Predictor selection aimed to use a parsimonious set of variables that captures the dominant sources of variation while avoiding collinearity-induced distortions in parameter estimates. Distance, velocity, and time-to-arrival are mathematically coupled through the kinematic identity $\mathrm{TTA}_0 = \frac{d_{on}}{v_{on}}$ and could therefore not be included simultaneously. Including all three predictors resulted in variance inflation and unstable coefficient estimates. The selected predictors were thus the experimentally manipulated variables oncoming vehicle velocity v_{on} and initial time-to-arrival TTA_0.

The third predictor was the visual looming rate τ^{-1}, which complements the kinematic variables by capturing non-linear visual expansion dynamics (Figure 1, bottom right panel). Lastly, decision outcome was included as a predictor in the RT analysis to account for systematic differences between *Drive* and *Wait* responses. This selection preserves coefficient interpretability, minimizes redundancy between physical and perceptual predictors, and aligns with the theoretical framing of the experiment.

2.3.2 Regression analysis formulations

To facilitate interpretation of regression coefficients and improve convergence, predictors were transformed following the recommendations of Schielzeth [24]. Continuous predictors were centred (.c) or standardised (.z), while categorical predictors were effect-coded (.e) as appropriate. In the RT analysis, decision outcome was effect-coded as $-1/1$ (dec.e), allowing both decision categories to contribute symmetrically to interaction terms. In contrast, for the decision outcome analysis, the binary response was coded as 0/1 to preserve its probabilistic interpretation [25].

The final regression formulations were:

$$\begin{aligned} RT \sim \quad & 1 + dec.e \times TTA_0.c + dec.e \times \tau^{-1}.z + \\ & dec.e \times v_{on}.e + \tau^{-1}.z \times TTA_0.c + \\ & \tau^{-1}.z \times v_{on}.e + TTA_0.c : v_{on}.e + (1|id) \end{aligned} \tag{1}$$

$$\begin{aligned} Decision \sim \quad & 1 + \tau^{-1}.z \times TTA_0.c + \tau^{-1}.z \times v_{on}.e + \\ & TTA_0.c : v_{on}.e + (1|id) \end{aligned} \tag{2}$$

where *dec* denotes the decision outcome (*Drive*/*Wait*) and $v_{on}.e$ the oncoming vehicle's velocity category (*Slow*/*Fast*).

2.3.3 Reference statistical method

The statistical analysis framework developed in this study builds on the methods introduced by Zgonnikov et al. [17], who analysed left-turn gap acceptance using both linear and logistic mixed-effects models (Appendix B).

2.4 Cognitive modelling of driver decisions

2.4.1 Drift diffusion model foundation

Decision behaviour was modelled using variants of the drift diffusion model (DDM) [8], which represents two-alternative choice as a stochastic accumulation process driven by task-relevant information. The DDM is used here to formalise how temporal, kinematic, and visual variables influence both decision outcome and decision timing within a single evidence-accumulation framework.

Formally, the rate of change of the decision variable is given by

$$\frac{dx}{dt} = s(t) + \varepsilon(t), \tag{3}$$

where $x(t)$ denotes the decision variable, $s(t)$ the drift term, and $\varepsilon(t)$ Gaussian noise. A response is registered when $x(t)$ reaches one of two decision boundaries, and total RT is given by the sum of the accumulation time and a non-decision time component.

In driver decision modelling, drift functions are typically specified either as functions of vehicle kinematics or as functions of visual-perceptual variables. The models studied here combine these two classes of drift formulation. Sections 2.4.2–2.4.6 describe the model components used in this study, and Sections 2.4.7–2.4.9 define the specific model variants and fitting procedure.

2.4.2 Drift diffusion model design

The classic drift diffusion model [8], [26] provides a flexible foundation for extension of model components, as all components can be represented or replaced with new, more complex functions. The four key components that can be designed are the starting bias x_0 (the starting value of the decision variable), the drift function $s(t)$ (the rate of change of the decision variable dependent on task-relevant variables), the decision boundary $B(t)$ (the value at which the decision variable is of sufficient magnitude to make a decision) and the non-decision time (the added time due to perceptual and motor delays). Additional constraints can be added, but for the application in driver decision modelling, these four components suffice.

In this study, the aim is to find a model formulation with kinematic variables that successfully emulates the behaviour of participant populations, while being robust to variability in training data and conceptually generalisable across different task paradigms. This goal is taken into account by adhering to three design guidelines: 1) minimise the number of variables; 2) minimise the task-specificity of model components; 3) maximise robustness to population noise. Below, the design of every model component is explained, in chronological order according to the assumed cognitive processes.

2.4.3 Starting bias of the decision variable

The bias x_0 of the model describes the offset towards one of the two choice alternatives at the start of the accumulation process (Figure 1, top-left). Conceptually, this represents a belief or preference for one outcome, either due to individual differences or a learned behaviour. A bias of zero thus means that there is no preference for

any choice alternative, while a non-zero bias signifies a preference for one choice. The bias can be formulated as any function of variables describing the initial state of the perceptible environment, as a single free parameter, or simply as zero (assuming no bias). In this study, only the latter two approaches were tested. This adheres to design guideline 1, but is also supported by recent work in driver decision modelling showing that using complex bias definitions doesn't necessarily improve model performance [17], [18].

2.4.4 Non-decision time due to neurological delays

The non-decision time describes the time added by delays in perceptual and motor-related neurological processes. These delays occur due to dynamics in electrical signal propagation in neurons and muscles. This non-decision time X_{NDT} varies across individuals and trials and follows a normal distribution with characteristic mean μ_{NDT} and standard deviation σ_{NDT}, thereby adhering to design guideline 3:

$$X_{NDT} \sim \mathcal{N}(\mu_{NDT}, \sigma^2_{NDT}) \tag{4}$$

2.4.5 Drift function formulation

The drift ($s(t)$ for eq. 3) describes the direction of change of the decision variable $x(t)$, indicated by arrows in the schematic in Figure 1, that is attributed to task-related variables (as opposed to the added Gaussian noise $\varepsilon(t)$). Conceptually, the drift function describes the relative influence of the different types of perceivable information on the belief that a certain decision (one of the decision boundaries) is correct. In this study, the drift definition is the main source of model variation. The variables of the drift function are zero-order (describing positions and dimensions) or first-order (describing rates of change of positions and dimensions). All variations of drift definition can be found in Appendix D, but a prevalent foundation for drift functions across gap-acceptance literature is provided below [11], [12], [16].

$$s(t) = \alpha \left(TTA(t) + \beta_d d(t) - \theta_{crit} \right) \tag{5}$$

In this function, the time-to-arrival of an approaching vehicle is denoted by $\text{TTA}_o n(t) = \frac{d(t)}{v(t)}$. Here, $d(t)$ and $v(t)$ are the distance and velocity of the approaching vehicle at time t. α, β_d, and θ_{crit} are free parameters, where α quantifies the contribution of TTA_{on}, β_d quantifies the contribution of distance, and θ_{crit} represents an arbitrary critical value to which the other contributions are compared. The complete set of variables that will be used to define drift functions in this study further includes: velocity $v(t)$, visual angle $\theta_v(t)$, and visual looming $\tau^{-1}(t)$ with their respective parameters β_v, β_θ, and β_τ. This set of variables adheres to design guideline 2.

2.4.6 Decision boundary

The decision boundary $\pm B$ describes the values of the decision variable at which a choice is selected, corresponding to the upper and lower thresholds shown in Figure 1. Conceptually, this means that the boundary describes the amount of evidence an individual requires before believing a certain choice is correct. As indicated by the $\pm$ sign, the decision boundary value is used for both the positive and negative boundaries. When the decision variable crosses the positive boundary, the outcome is to *Drive*, meaning to advance through the narrow passage. Conversely, crossing the negative boundary results in the choice to *Wait* for the oncoming vehicle to pass before advancing through the narrow passage. In literature, common definitions are: a fitted constant B, or a function that collapses with increasing TTA [12], [17], [18]:

$$B = \pm \frac{b_0}{1 + e^{-k(TTA(t) - \tau)}} \tag{6}$$

where b_0 is the initial boundary, k quantifies the influence of TTA on the rate of collapse, and τ represents some critical TTA value. The inclusion of a collapsing boundary is essential for good model fit in some driving task paradigms [11]. However, in this study, only exploratory models included this type of boundary, and the focus models did not require a collapsing boundary to accurately describe decision patterns.

2.4.7 Reference Drift Diffusion Models

The goal for DDM design in this study was to establish a unified and theoretically grounded starting point for generalisable kinematics-based DDMs. Therefore, the framework developed in the present study builds on methods from Zgonnikov et al. [17] and Theisen et al. [18]. The features of their approach suitable for cross-task generalisation are: they assume a constant decision boundary, a single free bias parameter, and a normally distributed non-decision time. They constitute the "base models" of the formulations presented in section 2.4.8. Detailed model definitions and results can be found in Appendix B.

2.4.8 Six longitudinal kinematics DDMs for the narrow passage study

This paper focuses on six longitudinal kinematics (LK) DDM variants that were selected based on the reference models from the literature presented in the previous section [17], [18], additional modelling approaches from emergency manoeuvring studies [27], [28], results of the statistical analyses in Section 2.3, and the qualitative performance of the models during exploratory modelling across the lab and online datasets. The six DDMs vary only in their drift function, and vary across two axes: **1)** the base kinematic variable (distance denoted by DLK or velocity denoted by VLK); **2)** the optional inclusion of a visual variable (visual angle θ, or looming τ).

Table 3: Naming convention for six DDMs varied across two conceptual axes: base kinematic variable, and visual variable.

Class	**Base**	$\boldsymbol{\theta_v}$**-extension**	$\boldsymbol{\tau^{-1}}$**-extension**
Distance-based	DLK	DLK-θ	DLK-τ
Velocity-based	VLK	VLK-θ	VLK-τ

Conceptually, the result is a grid in which zero-order and first-order information is integrated across multiple information streams (Table 3). These streams provide visual, kinematic, and temporal information in variables

Table 4: Overview of the six focus model definitions.

Model	Drift function ($s(t)$)	Boundary	Bias	Non decision time	Predictor order	
					Kinematic	Visual
1: (DLK)	$\alpha\left(\mathrm{TTA}(t) + \beta_d d(t) - \theta_{crit}\right)$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	-
2: (VLK)	$\alpha\left(\mathrm{TTA}(t) + \beta_v v(t) - \theta_{crit}\right)$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1	-
3: (DLK-θ)	$\alpha\left(\mathrm{TTA}(t) + \beta_d d(t) + \beta_\theta \theta_v(t) - \theta_{crit}\right)$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	0
4: (VLK-θ)	$\alpha\left(\mathrm{TTA}(t) + \beta_v v(t) + \beta_\theta \theta_v(t) - \theta_{crit}\right)$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1	0
5: (DLK-τ)	$\alpha\left(\mathrm{TTA}(t) + \beta_d d(t) + \beta_\tau \tau^{-1}(t) - \theta_{crit}\right)$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	1
6: (VLK-τ)	$\alpha\left(\mathrm{TTA}(t) + \beta_v v(t) + \beta_\tau \tau^{-1}(t) - \theta_{crit}\right)$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1	1

that require different levels of cognitive transformation to be *perceived*. For instance, the visual angle is a direct signal strength from the retina, but the brain must perform geometric and temporal transformations to infer distances and velocities. The six focus models are all unique in their required perceptual transofrmations, as can be seen in Table 4

A broader set of eleven additional DDM variants was also evaluated, covering alternative combinations of drift components (e.g., looming-only, time-to-passing, visual-only designs), alternative boundary formulations (collapsing with TTA), and alternative bias definitions (distance- or visual angle-dependent). These models were included to probe the sensitivity of the LK framework and to assess whether other parametrisations varying in complexity offered significant improvements. However, most alternatives either produced unstable or weakly interpretable fits or offered no clear advantage over the selected formulations. For transparency, their definitions and comparative performance are summarised in Appendix D.

2.4.9 DDM modelling procedure

All model variants were implemented and fitted using the `pyDDM` framework [29]. Parameter estimation employed the library's built-in *differential evolution* optimiser, using the `auto` optimisation strategy and the Bayesian Information Criterion (BIC) as the objective function. Differential evolution provides a robust global search over the non-convex parameter space characteristic of drift diffusion models, enabling stable recovery of drift, boundary, and non-decision time parameters across model classes. Optimising BIC allows joint evaluation of predicted choice proportions and reaction-time distributions while appropriately penalising model complexity. Unless stated otherwise, optimiser hyperparameters were set to PyDDM's default values.

All models were fitted using the full participant sample within each dataset. This choice reflects the objective of the study: to infer a common decision mechanism at the population level rather than to predict individual driver behaviour.

2.5 Assessment of model robustness and generalisability

This subsection describes the procedures used to assess the robustness and generalisability of the proposed drift–diffusion models beyond their original training conditions. Robustness of the learned behaviour is evaluated across different sets of restricted training data. Generalisation of the learned behaviour is evaluated across both datasets, and generalisation of the model framework across a related driver decision-making task.

2.5.1 Condition-exclusion assessment (leave-one-condition-out)

The robustness of learned behaviour dynamics was assessed through condition-exclusion validation within each dataset. Models were trained on all but a subset of experimental conditions and evaluated on the held-out trials without refitting. Two exclusion strategies were used: (1) removing an entire TTA_0 level (i.e., both velocity conditions), and (2) removing a single TTA_0–velocity combination. These analyses test whether the drift formulation can predict behavioural structure in conditions that were not used during parameter estimation.

All six focus models were refitted for each exclusion scenario to assess how consistently different drift formulations captured behaviour under restricted training. Performance was evaluated using condition-level decision probabilities, mean RTs, and full reaction time probability density functions (RT-PDFs), allowing joint assessment of decision tendencies and distributional structure.

2.5.2 Cross-dataset assessment within the narrow passage task

Generalisation across datasets was evaluated by transferring models between the lab and online samples of the narrow passage task. Models were trained on one dataset and evaluated on the other without refitting. This analysis tests whether the inferred drift structure is robust to differences in experimental context, behavioural variability, and demographic characteristics, while holding the task structure fixed.

2.5.3 Cross-task assessment using the left-turn nudging experiment

To assess whether the inferred drift structure generalises beyond the narrow passage paradigm, LK-based drift diffusion models were additionally evaluated on data from the left-turn gap-acceptance experiment, a related but distinct decision-making task. Unlike the narrow passage task, participants in the left-turn experiment had continuous control over steering, braking, and acceleration, and interacted within a higher-fidelity simulated environment.

For this analysis, the base DLK and VLK drift formu-

lations were refitted to the nudging dataset using the same modelling framework and estimation procedure as in the narrow passage task. visual drift extensions were not considered for this dataset, as video or other visual information from the experimental setup and trials was not readily available. Model performance was evaluated by comparing predicted decision proportions and reaction-time structure across nudging conditions. This validation tests whether the core evidence-accumulation dynamics captured by the LK formulations are specific to the narrow passage scenario or reflect a more general decision mechanism applicable across interaction tasks.

3 Analysis of experimental results

This section presents the full descriptive analysis of the experimental results from both datasets. Figures 2 and 3 provide a visual representation of the patterns in decision outcome and RT with respect to TTA_0 and v_{on}, and Tables 5-6 provide quantitative support to the observed patterns. Figure 2 summarises the overall RT distributions in the lab and online datasets, providing a compact view of between-dataset differences in response timing. Secondly, Figure 3 relates both decision outcome and RT to the independent variables initial time-to-arrival (TTA_0) and oncoming vehicle velocity v_{on}. Finally, Tables 5 and 6 report the mixed-effects regression results for the decision outcome and RT. In the following, behavioural patterns observed in the lab and online are described and supplemented with their statistical support. More detailed analysis of the regression analyses can be found in Appendix C.

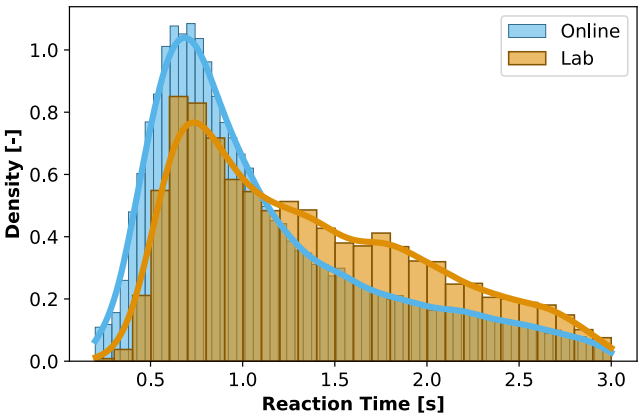


Figure 2: RT density of both datasets. Bin size was chosen automatically by the Seaborn library for Python. Lines represent the kernel density estimates of the histograms. Adapted from Miller et al. [15]

3.1 Separate behaviour patterns for lab and online data

The lab and online datasets are considered separately because they differ in both demographic characteristics and the degree of experimental control. While the lab sessions ensured consistent environmental conditions, the online study presumably contains higher variability in display characteristics, attention, and experimental context.

Figure 2 shows the RT distributions for both datasets. The online data exhibit a clear right skew, consistent with shorter mean RTs (M=1.11s, SD=0.62s) compared to the lab sample (M=1.36s, SD=0.65s). Both distributions appear broadly unimodal, though the distribution for the lab data is less smooth, likely due to the smaller sample size. The datasets also differ in overall decision proportion P($Drive$): the lab sample shows P($Drive$)=0.515, whereas the online sample shows P($Drive$)=0.462. The online participants, therefore, appear more conservative. This relationship is especially seen in trials with long TTA_0, where online participants were less likely to choose to drive, as can be seen in Figure 3.a.

To evaluate whether demographic factors could account for the observed differences between the lab and online datasets, exploratory analyses were conducted. Because differences in experimental control limit the interpretability of direct cross-dataset comparisons, demographic effects were primarily assessed within each dataset. Within-dataset correlations between demographic variables (age, driving experience, habitual kilometres) and both RT and decision outcome were negligible. An exploratory comparison of participant-level averages across datasets suggested a possible age-related trend; however, given the reduced experimental control in the online setting, this pattern is interpreted cautiously and not considered evidence of systematic demographic influences. Additional details are provided in Appendix E.

3.2 Patterns in decision outcome

Interpretation of the descriptive plot in Figure 3.a and the results of mixed-effects logistic regression analysis revealed some key insights about the dynamics of the decision outcome of the narrow passage task. Firstly, decision outcomes in the narrow passage task are clearly shaped by strong and consistent TTA and v_{on} effects that are directly distinguishable in the experimental data. In contrast, the effects of visual predictors and joint effects between predictors are revealed primarily through statistical modelling and act as modulators of behaviour rather than primary determinants.

3.2.1 Drive probability strictly increases with time-to-arrival TTA_0

Figure 3.a shows that, in both datasets, the probability of choosing *Drive* always increases with initial time-to-arrival (TTA_0). This pattern is consistent across velocity conditions. The regression results confirm this pattern in both the lab and online samples (Table 5).

3.2.2 Drive probability increases with oncoming vehicle velocity v_{on}

A similarly consistent pattern is observed for oncoming vehicle velocity v_{on}. As shown in Figure 3.a, a higher v_{on} systematically results in a higher *Drive* probability in both datasets and across TTA_0 conditions. This effect is strongly supported by the regression results, which show a large effect size between v_{on} and the *Drive* probability in both the lab and online data.

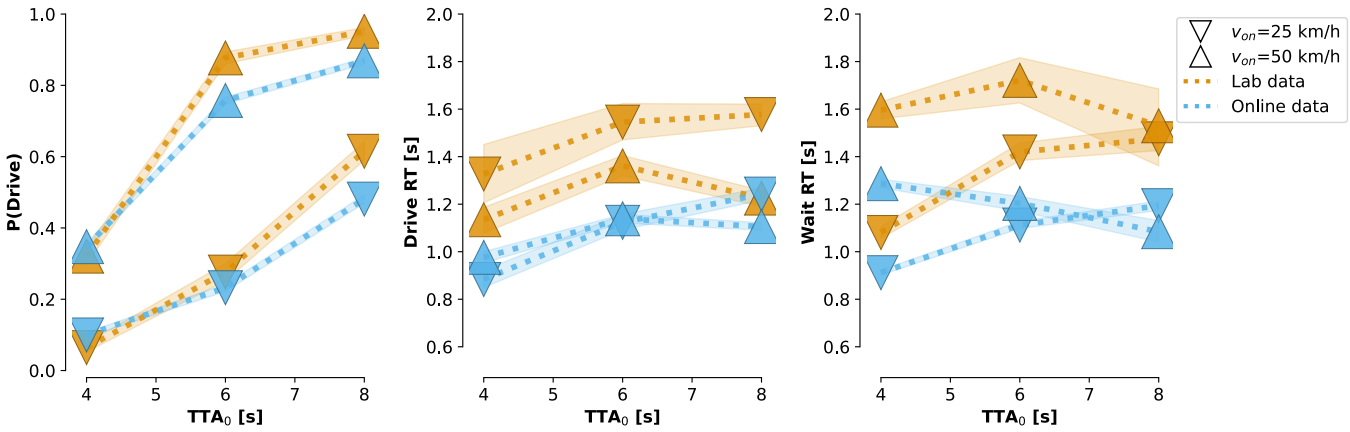


Figure 3: Effect of initial time-to-arrival TTA_0 and oncoming vehicle velocity v_{on} the probability of reaching a *Drive* decision and on the RT of a *Drive* or *Wait* decision. Lines are for clarity, but represent no data. The 95% confidence intervals for the population means are indicated by shading.

Table 5: Mixed-effects logistic regression results for decision outcome in the lab and online datasets. Fixed-effect coefficients reflect the contribution of each predictor to the log-odds of choosing *Drive*, conditional on all other predictors and interactions included in the model. Continuous predictors were centred (.c) and, where indicated, standardised (.z). Categorical predictors were effect-coded (.e).

	Lab dataset				**Online dataset**			
Predictor	Est.	SE	z	p	Est.	SE	z	p
(Intercept)	0.274	0.179	1.532	0.126	-0.144	0.107	-1.341	0.180
$\tau^{-1}.z$	-0.053	0.067	-0.790	0.430	-0.265	0.032	-8.268	***
TTA_0.c	0.929	0.033	28.12	***	0.686	0.013	50.87	***
v_{on}	1.696	0.047	36.18	***	1.306	0.018	72.01	***
$\tau^{-1}.z : TTA_0$.c	0.143	0.028	5.171	***	0.176	0.012	15.23	***
$\tau^{-1}.z : v_{on}$	-0.586	0.049	-12.08	***	-0.087	0.025	-3.518	***
TTA_0.c : v_{on}	-0.076	0.030	-2.510	*	0.039	0.012	3.266	**

Significance codes: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

3.2.3 Visual looming rate consistently modulates the effects of TTA_0 and v_{on} on Drive probability across both the lab and online datasets.

Effects related to visual looming rate (τ^{-1}) are not directly apparent from the descriptive plots. However, the regression results reveal that looming contributes to the decision outcome primarily through the modulation of kinematic effects. Across datasets, the main effect of the standardised looming rate ($\tau^{-1}.z$) is not consistent. In Table 5, the interaction effects that include looming are statistically significant and consistent in sign. Specifically, the positive interaction between $\tau^{-1}.z$ and TTA_0 indicates that the effect of looming depends on the available time budget at onset, while the negative interaction between $\tau^{-1}.z$ and velocity shows that looming becomes less influential at higher speeds.

3.2.4 Interaction effects of kinematic predictors are inconsistent across both the lab and online datasets.

Finally, the interaction between TTA_0 and velocity is difficult to infer from the descriptive plots in Figure 3.a. The regression results are also inconsistent in sign across datasets. This indicates that, although the main kinematic effects are consistent across experimental settings, their joint influence on decision outcome differs between the lab and online samples.

3.3 Patterns in reaction time

Interpretation of the patterns in RT associated with temporal, kinematic, visual, or decisional predictors proved to be much less straightforward than when considering the decision outcome. Figure 3.b-c shows that the influences of independent variables TTA_0 and v_{on} do not follow strict patterns across conditions or datasets. Nonetheless, as explained below, three key insights were revealed by the results of the linear regression (Table 6). Overall, RTs in the narrow passage task depend on complex interactions of temporal, kinematic, visual, and decisional factors. While all main effects are statistically significant and of notable effect size, they cannot be interpreted independently due to the magnitude and consistency of interaction effects. The small influence of decision outcome on RT is not consistently observable across experimental conditions and is therefore not readily interpretable as a dominant predictor of response timing.

Table 6: Linear mixed-effects model results for RTs in the lab and online datasets. Type-III ANOVA F statistics reflect each predictor's contribution to variance in log-RT, conditional on all other model terms. Continuous predictors were centred (.c) or standardised (.z), and categorical predictors were effect-coded (.e). Full tables are reported in Appendix C.

	Lab dataset				**Online dataset**			
	LMM coefficients		**ANOVA-III**		**LMM coefficients**		**ANOVA-III**	
Effect	Est.	SE	F	p	Est.	SE	F	p
(Intercept)	1.30	0.02	–	–	1.22	0.01	–	–
decision.e	-0.05	0.00	195.8	***	-0.01	0.00	83.06	***
TTA_0.c	0.31	0.00	19861	***	0.29	0.00	79170	***
$\tau^{-1}.z$	0.78	0.00	32651	***	0.77	0.00	122898	***
v_{on}.e	0.26	0.00	5955	***	0.24	0.00	30048	***
decision.e:TTA_0.c	0.04	0.00	209.6	***	0.03	0.00	861.4	***
decision.e:$\tau^{-1}.z$	0.06	0.00	259.1	***	0.03	0.00	400.6	***
decision.e:v_{on}.e	-0.01	0.00	9.610	**	-0.00	0.00	3.630	0.057
TTA_0.c:$\tau^{-1}.z$	0.11	0.00	3175	***	0.10	0.00	11868	***
$\tau^{-1}.z$:v_{on}.e	-0.09	0.00	740.8	***	-0.08	0.00	2226	***
TTA_0.c:v_{on}.e	0.06	0.00	511.0	***	0.07	0.00	5471	***

ANOVA significance codes: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Intercepts are not tested in the Type-III ANOVA.

3.3.1 Reaction time generally increases when more time is available

Across both datasets, RT generally increases with the initial time-to-arrival (TTA_0) at decision onset, particularly at low oncoming vehicle velocity v_{on} (Figure 3.b–c). This indicates that participants respond more slowly when more time is available to resolve the interaction. At high velocity, this relationship changes across TTA_0 conditions, decision outcomes, and datasets. This indicates that the influence of TTA_0 on response timing depends on the kinematic context and, in some cases, on the selected action. The regression analysis supports this interpretation by showing a relatively large effect of TTA_0 alongside statistically significant interactions with other predictors (Table 6). Together, these results indicate that a larger time budget generally prolongs response timing, but this effect is modulated by kinematic, visual, and decision context.

3.3.2 Individual predictors show non-uniform effects on reaction time due to strong interaction effects

Except for the general TTA_0 trend described above, none of the predictors (v_{on}, τ^{-1}, and decision outcome) show consistent effect across all other conditions or across both online and lab data. The regression analysis clarifies this ambiguity by showing that the effects on RT of all predictors and all interactions except *v_{on}:decision* are statistically significant. The effect sizes of the temporal, kinematic, and visual predictors are all large, and their paired interactions are similar. Notably, the effect size of τ^{-1} is largest, which could be a consequence of standardisation. As a result, the visible effects of individual predictors depend critically on the configuration of other variables, explaining why main effects are difficult to interpret without considering the interaction structure.

3.3.3 Decision outcome weakly modulates reaction time

The decision outcome has a modest but statistically significant influence on RT. Across both datasets, *Drive* responses are, on average, slightly faster than *Wait* responses, as indicated by small negative coefficient estimates in the regression analysis (Table 6). This effect is not consistently visible across TTA_0 and velocity conditions in Figure 3. Most interaction effects involving decision outcome are statistically significant, although their magnitudes are relatively small. In the lab results, a pattern is visible where *Drive* RTs seem negatively related to v_{on}, and *Wait* RTs are positively related to v_{on}, but this pattern does not hold in the online data. These results indicate that the influence of decision outcome on RT represents a secondary rather than a dominant determinant of response timing.

3.4 Summary of experimental results analysis

Inspection of the experimental results revealed that decision behaviour in the narrow passage task is governed by strong and reliable temporal and kinematic constraints, with visual information contributing in a systematic but condition-dependent manner. Across both datasets, the patterns observed in decision outcomes are highly similar, despite the differences in overall RTs. This indicates comparable underlying decision mechanisms operating under different levels of experimental control. Decision outcomes are largely determined by initial time-to-arrival and oncoming vehicle velocity, whereas visual looming primarily modulates these kinematic influences rather than acting as an independent driver of choice. RTs, in contrast, exhibit a more complex structure: while more available time generally leads to slower reactions, especially at low velocities, the effects of individual predictors are highly non-uniform. RT patterns are largely determined by interactions between temporal, kinematic, visual, and decisional factors. These findings indicate that observed behaviour cannot be fully explained by

isolated stimulus–response mappings or main effects alone. Instead, they point to an underlying decision process in which multiple sources of information are integrated over time to determine both choice and response timing. These findings motivate the use of the drift diffusion modelling framework in the following section, which is employed to model the cognitive process by which visual, kinematic, and temporal information is integrated over time.

4 Results of Drift Diffusion Modelling

This section reports the results of applying the DDM framework to decision behaviour in the narrow passage task. The analysis focuses on the six longitudinal kinematics (LK) DDM variants from Section 2.4.8, spanning distance- and velocity-based drift formulations, referred to as DLK and VLK models, respectively, with and without visual augmentation. Tables 7 and 8 and Figure 4 summarise model fit quality, best-fitting parameter estimates, and predicted decision characteristics across lab and online datasets. All models in this section were trained on the complete datasets, without excluding participants or experimental conditions, to provide the strongest basis for capturing the interactions between temporal, kinematic, and visual variables identified in the statistical analyses. Fit performance of additional exploratory models is reported in Appendix D. To aid interpretation of the dense set of results that follows, the main empirical findings from the DDM analyses are summarised below before being examined in detail.

Table 7: Model fit comparison of all six focus models trained on the labo and online datasets. All models include the same temporal variable TTA(t), which is not visible in the table. Ranking is relative within each dataset. The best performance is attained by the VLK-τ model in the laboratory dataset (green) and the VLK model in the online dataset (blue).

Setting	Model	BIC	Rank	κ	Kin. var.	Vis. var.
Lab	DLK	23269	5	7	d	None
	VLK	23106	3	7	v	None
	DLK-θ	23233	6	8	d	θ_v
	VLK-θ	23102	4	8	v	θ_v
	DLK-τ	23009	2	8	d	τ^{-1}
	VLK-τ	23000	1	8	v	τ^{-1}
Online	DLK	102269	6	7	d	None
	VLK	101625	1	7	v	None
	DLK-θ	102327	5	8	d	θ_v
	VLK-θ	101630	2	8	v	θ_v
	DLK-τ	101927	4	8	d	τ^{-1}
	VLK-τ	101645	3	8	v	τ^{-1}

Fit performance of all exploratory models reported in Appendix D

Table 8: Estimated parameter values for the best-performing models in each dataset.

Model	α	β_v	β_τ	θ_{crit}	B	x_0	μ_{ndt}	σ_{ndt}
VLK-τ								
Lab	0.433	0.394	5.260	10.419	1.092	0.060	0.374	0.095
VLK								
Online	0.296	0.502	—	11.027	0.949	-0.003	0.302	0.112

Parameters of all exploratory models reported in Appendix D

4.1 Key findings from drift diffusion modelling

- All studied DDMs achieve competitive fits to both lab and online data of the narrow passage task.
- Fitted DDMs capture comparable decision and timing structure across lab and online datasets within the narrow passage paradigm, despite differences in sample size and experimental control.
- Velocity-based drift formulations show a modest but consistent advantage in overall fit quality compared to distance-based formulations.
- Augmenting the drift with visual looming yields selective and condition-dependent improvements, most prominently for distance-based DDMs.
- Across model variants, DDMs do not reliably capture the shift in timing structure, and consequently underestimate the associated RTs.

Before presenting the detailed results, Section 4.2 compares the six focus models representing two model classes, as described in Section 2.4.8, by their fit performance. Section 4.3 details the mean decision probability and RT predictions from the DLK and VLK model classes. Sections 4.4-4.5 then further explain those results by analysing the full RT-PDFs per experimental condition. Finally, Section 4.6 sheds light on the influence of including visual variables in the drift function.

4.2 Longitudinal kinematics DDMs achieve robust and comparable fits across lab and online datasets

Table 7 compares the fit performance of all six LK-based DDMs across the lab and online datasets. Before interpreting differences in model ranking, it is important to clarify that BIC differences smaller than approximately 10 are of the same order of magnitude as variations arising from sensitivity to local optima due to different initial parameters of the optimisation algorithm, and are therefore not reliably interpretable. Consequently, only larger BIC separations are taken to indicate meaningful differences in fit quality, whereas small rank changes between closely performing models are interpreted with caution. All models consist of seven or eight free parameters, making them dimensionally similar, while the substantially larger size of the online dataset explains the overall shift toward larger absolute BIC values compared to the lab data.

4.2.1 Distance- and velocity-based models show similarly high fit performance

Under this interpretation, the BIC patterns reveal a clear and consistent structure across datasets. Across five out of six direct comparisons, velocity-based drift formulations (VLK) outperform their distance-based counterparts (DLK) by BIC margins that substantially exceed BIC thresholds associated with optimisation variability. DLK and VLK formulations only attain comparable BIC values for the lab dataset in looming-augmented form. This pattern indicates a robust and systematic advantage for velocity-based drift formulations within the LK framework across experimental contexts.

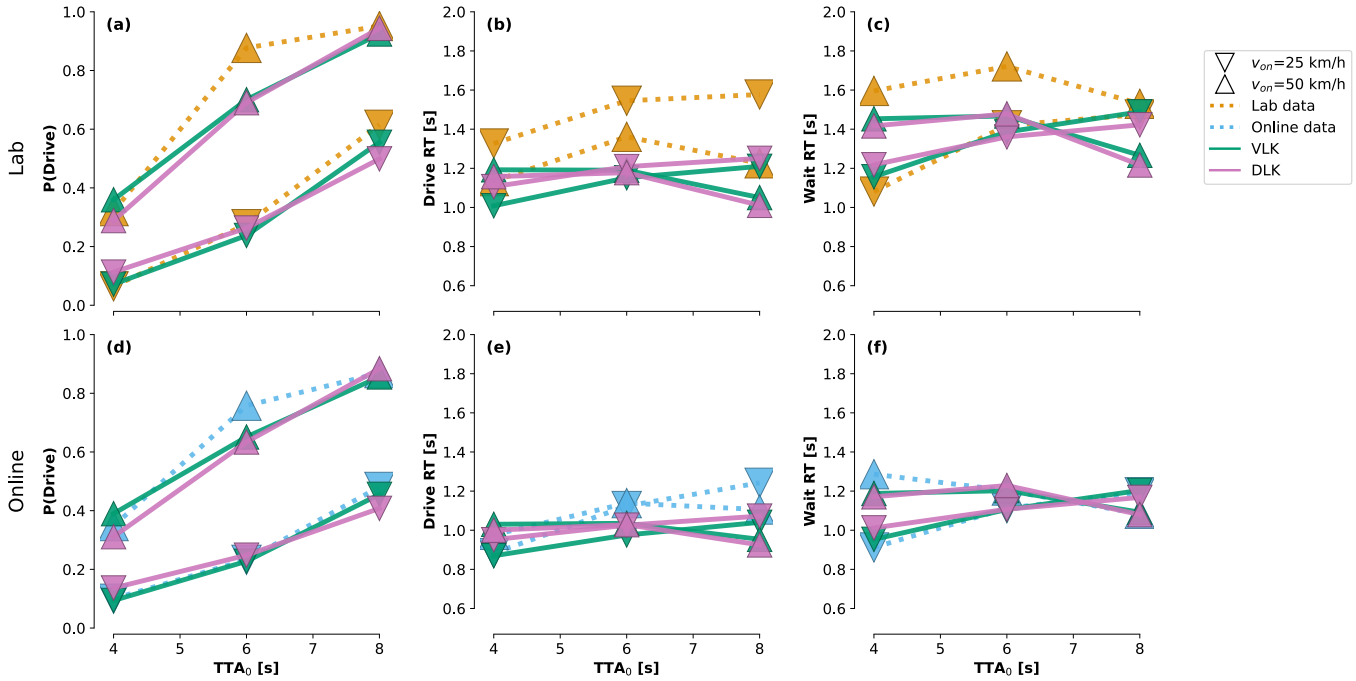


Figure 4: Population-averaged *Drive* probability and mean RTs from measured data and predicted by the distance- and velocity-based base models for every experimental condition. 95% confidence intervals are omitted for clarity.

4.2.2 Fitted model parameters are consistent across lab and online datasets

Parameter estimates are stable across different drift formulations. In Table 8, the model parameters for non-drift components are very comparable. This indicates that the models achieve a strong foundation of parameters that leaves space for variation only in drift-related parameters. Models fitted separately to the lab and online datasets also yield highly similar drift-related parameter estimates across both LK-based DDM classes. This consistency indicates that the fitted drift parameters capture comparable decision dynamics despite substantial differences in sample size, participant population, and experimental control. For brevity, the full set of fitted parameters for each model is not reported here, but is provided in Appendix D.

4.2.3 Visual extensions of the drift function selectively improve fit across both datasets and kinematics-based model classes

The contribution of visual augmentation is more nuanced. In the lab dataset, models augmented with visual looming tend to rank slightly above their kinematic-only counterparts, particularly for distance-based formulations. In the online dataset, however, the velocity-only LK model attains the lowest BIC, with visual-augmented variants ranking comparably but slightly higher. This suggests that while incorporating first-order visual information can yield modest improvements, such benefits are sensitive to dataset characteristics and are secondary to the dominant distinction between velocity- and distance-based drift formulations.

In conclusion, two general observations follow from these results. First, although velocity-based formulations consistently outperform distance-based formulations, all LK-based DDM variants achieve competitive fits and capture broadly similar behavioural structure. Second, looming augmentation introduces context-dependent improvements that are modest relative to the strong and consistent effect of the kinematic drift formulation itself. Visual angle augmentation has a mostly negligible effect. These results motivate a closer examination of model predictions in the following sections, where differences between drift formulations and visual augmentations are analysed in more detail.

4.3 Studied models capture decision tendencies but underestimate delayed reaction times

To illustrate how the drift formulation shapes model predictions, this section compares the *Drive* proportions P(*Drive*) and mean RTs generated by the two baseline LK-based DDM classes (DLK and VLK) as they provide enough context to assess the primary differences in model performance. Within each model class, visual extensions are secondary in their effect on predictive performance, so their effects are discussed separately in Section 4.6. Figure 4 shows the predicted mean P(*Drive*) and mean RTs from the lab-trained DLK and VLK models across all TTA$_0$–v_{on} conditions, alongside the measured lab data, as they provide the most controlled context.

4.3.1 Drive probability is captured robustly by distance- and velocity-based DDMs

For the decision outcome, both DLK and VLK models closely reproduce the systematic increase of P(*Drive*) with TTA$_0$ and the positive influence of v_{on} across both datasets. Both models also align closely with the absolute measured probabilities across most conditions. Differences between the two formulations are small and, at the level of mean P(*Drive*), only occur in specific conditions, but not systematically. Some challenging regions remain in the condition space of the data, such as the steep transition

between TTA_0=4 and TTA_0=6 at high velocity, and the TTA_0=8, v_{on}=25 condition. The difficulty that both models have in these conditions indicates that interactions of variables remain challenging to capture for both model classes.

4.3.2 Mean RT predictions from distance- and velocity-based models capture conditional patterns across lab and online datasets

The model fit for mean RTs is good in the sense of reproducing patterns, but less so in reproducing absolutely measured averages. For RTs, both DLK and VLK mimic the main patterns visible in the measured data: at low velocity, predicted *Drive* RTs increase with TTA_0, while at high velocity the relationship becomes weaker and, in some conditions, non-uniform. However, both models have difficulty reproducing the absolute mean RTs, especially in low-v_{on} *Drive* decisions and high-v_{on} *Wait* decisions. This effect is most distinct in the lab-trained models. Another notable pattern is best observed in Figure 4.e-f: Both DLK and VLK models worsen in their RT prediction (underestimating RTs) with increasing TT_0 for *Drive* decisions and with decreasing TTA_0 for *Wait* decisions.

While the DLK and VLK models reproduce the conditional patterns in mean RTs, the origin of systematic underestimation observed in several regions of the condition space cannot be fully understood from averages alone. Averaging RTs per experimental condition can obscure informative temporal structure in decision behaviour. Such a structure is captured in the full reaction-time probability density functions (RT-PDFs), whose peak height, skewness, and tail mass directly reflect properties of the underlying diffusion process. To disentangle the combined influence of these distributional features, the following section examines the fit of predicted RT-PDFs for each experimental condition and decision outcome.

4.4 RT distributions are well predicted and explain the qualitative fit of the studied models

To examine the temporal structure present in the RT data, Figure 5 shows the full RT-PDFs for the lab dataset, overlaid with predictions from the DLK and VLK base models. The RT-PDFs make it possible to assess both the structure of the measured response distributions and how closely each model class reproduces them.

4.4.1 Measured RT distributions reveal interpretable but uneven patterns over time

Across most experimental conditions, the measured RT-PDFs exhibit clear and systematic structure. *Drive* decisions are typically characterised by right-skewed distributions, with a rapid rise to a single peak followed by a long right tail, whereas *Wait* decisions tend to show broader distributions with lower peaks and more gradual decay. This structure is most apparent in conditions with a large number of observations, where participants' decision timing follows consistent and repeatable patterns. In contrast, measured RT-PDFs become irregular or weakly structured in data-sparse decision scenarios. In these cases, the broad confidence intervals observed in Figure 3 can be directly traced back to variability and sparsity in the underlying RT distributions.

4.4.2 Ambiguous conditions exhibit a delayed decision structure

Beyond data sparsity, several conditions exhibit RT-PDFs that are behaviourally ambiguous despite moderate sample sizes. In such cases, the measured distributions are less skewed and noticeably flatter than the canonical right-skewed shape, as observed for *Drive* at (TTA_0=8, v_{on}=25) and *Wait* at (TTA_0=4, v_{on}=50). Notably, these conditions are equal in initial distance to the oncoming vehicle. These broader distributions likely reflect additional deliberation, during which participants postpone commitment while assessing the situation more carefully.

4.4.3 Predicted Drive and Wait decisions have distinctly different RT distribution shapes

Across conditions, both DLK and VLK predicted *Drive* and *Wait* RT-PDFs exhibit the same differences in peak height and tail shape as the lab data. Although the models capture this relative difference, they tend to underestimate the tail mass in the measured RT-PDFs. In addition, the models retain their canonical RT-PDF shapes even in conditions where the experimental data are sparse or kinematically ambiguous (i.e., the kinematic variables are not instantly interpretable). The skewed distribution has a lower mean RT than a more uniformly distributed one, which explains the underestimated mean RTs in the sparse and ambiguous conditions observed in Section 4.3.

4.4.4 Predicted RT distributions show no decisive advantage of distance- or velocity- based models

Overall, the RT-PDF analysis indicates that DLK and VLK models reproduce the dominant temporal structure of the measured data equally well. Differences between predicted RT-PDFs are minimal and restricted to slight variations in peak height. There is no clear pattern in the peak height differences of DLK and VLK models linked to decision outcome, TTA_0, or v_{on}.

In conclusion, RT-PDFs reveal strong overall performance in capturing the temporal structure of different decision scenarios by both DLK and VLK models. However, the lab data includes many conditions that exhibit irregular and flat RT distributions. Additionally, no clear cause can be identified for RT-PDF differences between DLK and VLK models. These findings motivate further inspection of online data, which improves the view of data-sparse decisions due to its larger sample size. To further investigate the ambiguous intermediate-distance conditions, the next section orders results by distance instead of along the experimental condition grid.

4.5 Relationships of decision timing with oncoming vehicle distance

To clarify the temporal patterns identified in the lab data, Figure 6 shows RT-PDFs from the online dataset reordered by initial distance, following from the kinematic

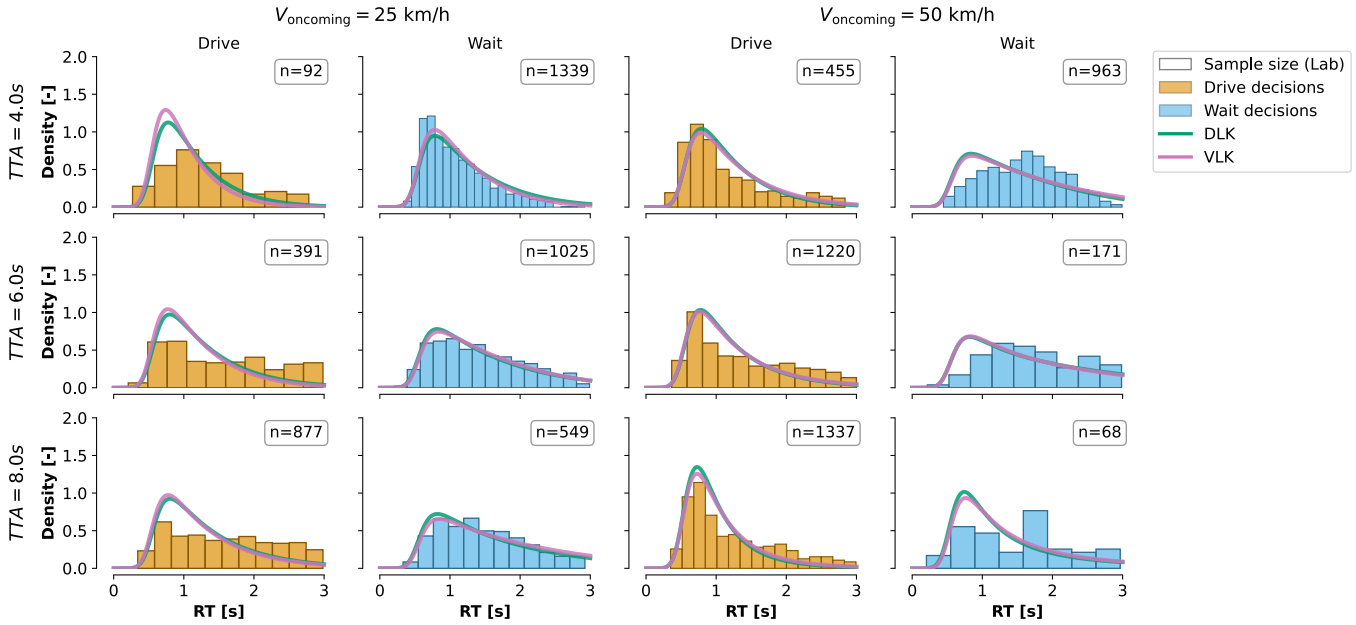


Figure 5: Reaction time probability density functions of the lab-trained base models overlaid on lab data, ordered by initial time-to-arrival on the vertical axis and by decision outcome and oncoming vehicle velocity on the horizontal axis. n denotes the number of trials in the lab dataset corresponding to each decision scenario.

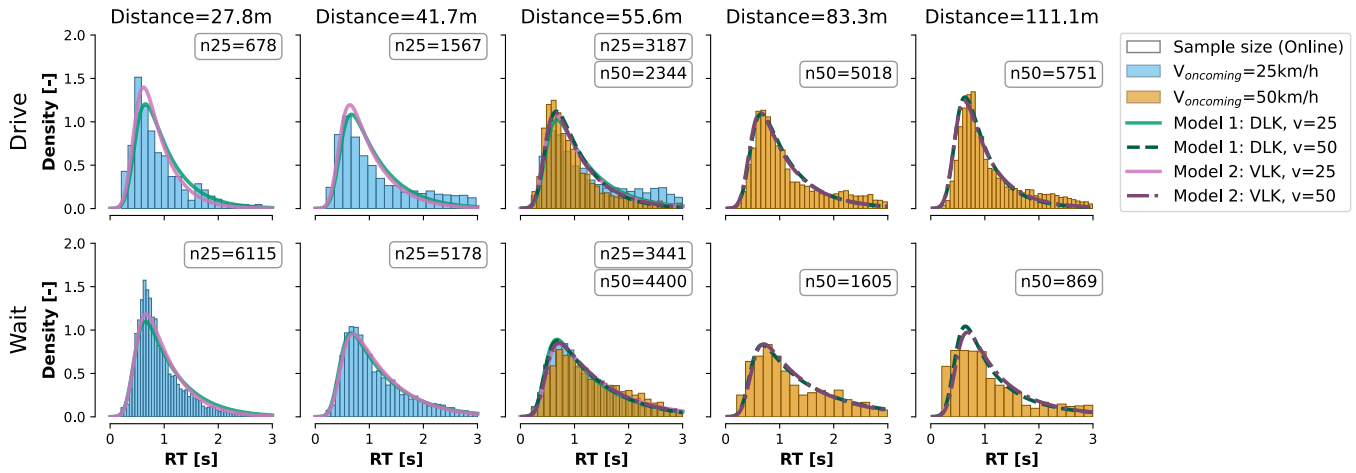


Figure 6: Reaction time probability density functions of the online-trained base models overlaid on online data, ordered by initial distance of the oncoming vehicle to the narrow passage on the horizontal axis and by decision outcome on the vertical axis. Colours indicate oncoming vehicle velocity. n denotes the number of trials contributing to each distribution and v_{on}.

identity TTA= $\frac{d}{v}$. The substantially larger sample size yields smoother empirical distributions and provides a clearer view of how RT structure evolves across distance and decision outcome.

4.5.1 Measured RT distribution show distance-dependent structure in decision timing consistently across lab and online datasets

Models trained on the lab dataset reproduce RT-PDF shapes that closely resemble those observed in the online data. Comparison of RT-PDFs in Figures 5 and 6 shows that lab-trained DLK and VLK models predict similar peak structure, skewness, and tail behaviour to models trained directly on the online dataset. This agreement suggests that the temporal dynamics learned from the lab data transfer robustly to the larger and less controlled online sample.

For *Drive* decisions, the RT-PDFs show a systematic progression with distance: sharp, high peaks at short distances, broader distributions at intermediate distances, and a partial return to sharper peaks at the longest distances. Across all conditions, however, the distribution tails remain relatively heavy and do not fully decay within 3 s.

Wait RT-PDFs exhibit a related but distinct pattern. Peak height decreases consistently with distance, without the recovery observed for *Drive* decisions, resulting in progressively broader distributions with more gradual decay. This divergence reinforces the asymmetry observed in the lab data, with *Wait* responses showing greater temporal variability across the condition space.

4.5.2 Decision outcome in delayed intermediate-distance decisions depends on TTA-decision interaction

The intermediate-distance condition contains overlapping TTA_0-v_{on} combinations and reveals a clear interaction between velocity, decision outcome, and timing. For *Wait* decisions, low-velocity trials produce sharper distributions, whereas high-velocity trials show broader RT-PDFs. The opposite pattern is observed for *Drive* decisions, where high-velocity trials dominate the even sharper distributions. This pattern indicates that decisions that were delayed are more likely to become a *Wait* decision if v_{on} is high (and the TTA_{on} is confirmed to be short), and more likely to become a *Drive* decision if v_{on} is low (and the TTA_{on} is confirmed to be long). This supports the complex interactional nature of response timing observed in Sections 3.3.2 and 4.3.

4.5.3 Distance- and velocity-based models capture distance-related trends

Both DLK and VLK models recover the dominant distance-dependent changes in peak height and overall distributional shape, with the VLK producing slightly higher and therefore better-fitting peaks in short-distance conditions. The advantage of VLK predictions over DLK predictions of RT-PDFs in short-distance conditions suggests that first-order kinematics (velocity) play a more important role in resolving these urgent conditions.

In summary, reordering the online RT-PDFs by initial distance reveals structured and repeatable relationships between decision timing, distance, velocity, and decision outcome that are only partially visible in the TTA_0-v_{on} ordered lab data. The larger online sample size reveals consistent distance-dependent changes in RT-PDF shape and clear timing differences between *Drive* and *Wait* decisions, as well as systematic effects of the TTA_0:decision interaction in delayed responses. Both DLK and VLK models reproduce the dominant distributional trends, but selectively underestimate tail mass, particularly for *Drive* decisions. Differences between DLK and VLK predictions are small and local, with a modest advantage for VLK in short-distance conditions.

4.6 Visual variables in the drift function provide local but asymmetric contributions to model fit

Reordering the mean decision proportions and RTs by initial distance reveals distance-dependent structure across conditions (Figure 7). Across all three measures, the two velocity conditions form distinct patterns within this distance-ordered organisation, indicating that both distance and v_{on} systematically shape the observed averages. The inclusion of visual looming modulates these patterns locally, but does not alter their dominant distance-dependent structure.

To assess the contribution of visual information in the drift function to model fits, Figure 7 compares the full predicted decision patterns of the base DLK with two visual-augmented variants: a zeroth-order visual angle term (DLK-θ) and a first-order looming term (DLK-τ). The DLK model is singled out because fit performance analysis in Section 4.2 revealed that the DLK model benefited most from the inclusion of visual variables. The influence of visual variables in VLK models is generally similar but much smaller compared to the effects described in the following. All analyses in this subsection use the online dataset, which provides smoother RT-PDFs and more complete coverage of delayed decision scenarios. The contribution of visual terms is examined separately at the level of decision proportions, mean RTs, and full reaction-time distributions.

4.6.1 Looming improves decision proportions without restructuring choice behaviour

Across experimental conditions, adding *visual looming* to the drift formulation improves P(*Drive*) predictions in the majority of conditions, including two out of three for low-velocity and high-velocity conditions. These improvements are small and do not alter the overall strength of pattern capture in decision behaviour by the base DLK model. In contrast, adding *visual angle* information has a negligible influence on decision proportions across the condition space. Together, these results indicate that first-order visual information can refine P(*Drive*) predictions of the DLK model, but does not fundamentally change the underlying decision dynamics.

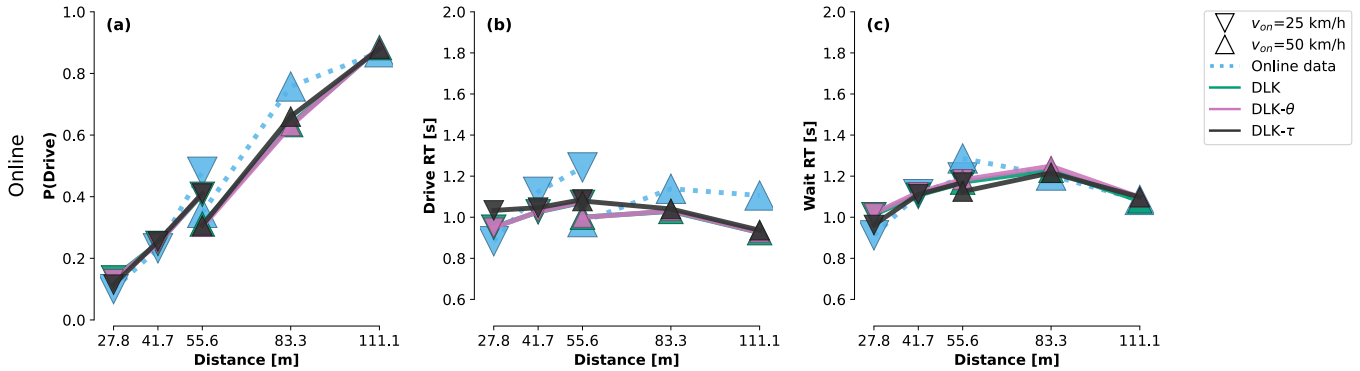
4.6.2 Visual looming introduces asymmetric trade-offs in reaction-time predictions

At the level of mean RTs, the effects of looming are strongly asymmetric across decision outcomes. For *Drive* decisions, looming degrades RT predictions across all conditions, increasing the discrepancy between predicted and measured mean RTs. For *Wait* decision, looming improves mean RT predictions only in the shortest-distance condition. Visual angle augmentation again has little systematic effect. These results reveal a clear trade-off: improvements in decision proportions and some *Wait* RTs come at the cost of systematically degraded temporal accuracy for *Drive* responses.

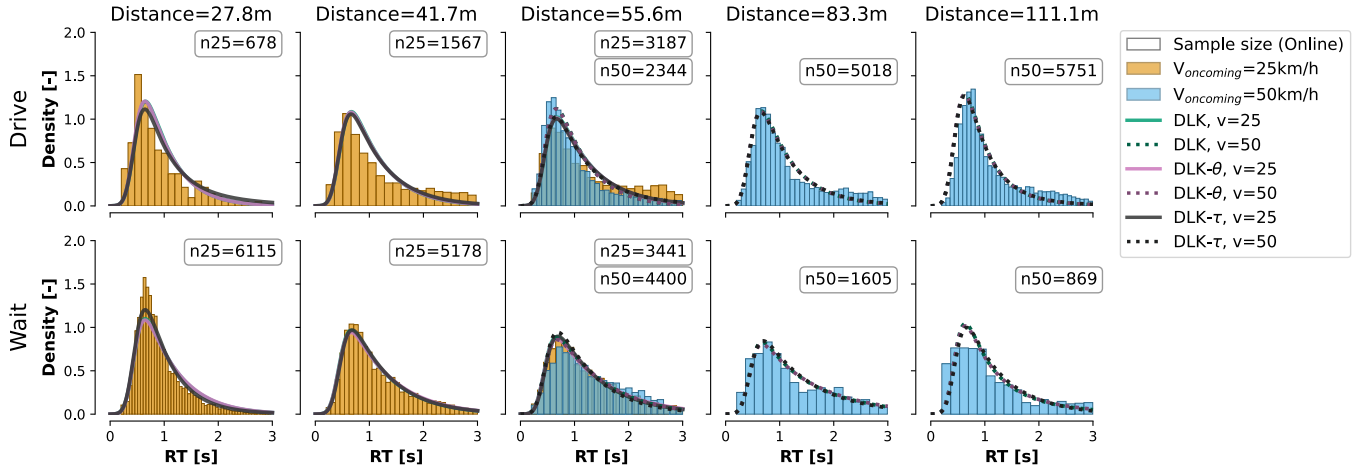
4.6.3 Looming selectively reshapes RT distributions in postponed decisions

Analysis of the full RT-PDFs shows that looming primarily affects slow v_{on} conditions and has a stronger influence on *Wait* than on *Drive* decisions. In low-velocity *Wait* conditions, looming improves alignment with the measured RT-PDFs, including in behaviourally ambiguous intermediate-distance conditions. For high-velocity *Wait* decisions, looming often produces distributions that are slightly too sharply peaked. In *Drive* decisions, looming has little effect overall, except in the shortest-distance condition, where it moderately worsens peak height, and in the intermediate-distance condition, where it underestimates peak height in the high-velocity condition. These selective effects indicate that visual motion cues modulate timing conditionally, rather than providing a consistent correction to delayed or urgent responses.

In conclusion, incorporating visual variables into the drift function yields systematic but asymmetric effects. Looming improves decision proportions in most conditions



(a) Population averaged *Drive* probability and mean RTs for the online-trained distance-based DDMs, overlaid on online data, ordered by initial distance between the oncoming vehicle and the narrow passage. 95 % confidence intervals omitted for clarity.



(b) Reaction time probability density functions of the online-trained distance-based DDMs overlaid on online data, ordered by initial distance of the oncoming vehicle to the narrow passage on the horizontal axis and by decision outcome on the vertical axis. Colours indicate oncoming vehicle velocity. n denotes the number of trials contributing to each distribution and v_{on}.

Figure 7: Comparison of predicted decision patterns by all three online-trained distance-based DDMS: DLK, DLK-θ, DLK-τ. Primarily, the inclusion of looming τ^{-1} in the drift function yields visible but asymmetric effects.

and selectively refines timing for low-velocity *Wait* responses, but simultaneously degrades *Drive* reaction-time predictions and fails to correct the broader underestimation of tail mass associated with postponed decisions. Overall, visual motion cues act as local modifiers of decision commitment and timing, rather than as global drivers of the accumulation dynamics captured by the base kinematic models.

4.7 Summary of drift diffusion modelling findings

Across lab and online datasets, the LK-based drift diffusion models provide a consistent account of decision behaviour in the narrow passage task. All DLK and VLK formulations achieve competitive BIC fits, reproduce the dominant patterns in P(*Drive*), mean RTs, and RT-PDFs. Velocity-based drift formulations show a modest but consistent advantage in fit quality, while distance-based formulations in particular benefit selectively from visual augmentation. Analysis of full RT-PDFs reveals that the models capture fast and common responses accurately, but systematically underestimate tail mass associated with postponed decisions, particularly in ambiguous conditions that require more temporal and kinematic resolution. Visual looming slightly refines decision proportions and selectively modulates timing in low-velocity *Wait* responses, but does not resolve these broader limitations. Together, these findings establish the strengths and challenges of the LK-based DDM framework within the narrow passage paradigm and motivate a dedicated validation analysis. The following section, therefore, evaluates model robustness and generalisability more explicitly using cross-dataset, cross-task, and cross-training-set validation procedures.

5 Generalisation and robustness of studied decision models

While the preceding sections established that longitudinal kinematics-based drift diffusion models (LK-DDMs) capture decision behaviour well within the narrow passage task, a central goal of this work is to demonstrate that the predicted patterns from trained DDMs transfer across datasets of varying size, experimental context, and coverage of the condition space, and across decision tasks. This section evaluates the robustness and generalisability of the studied LK-DDMs beyond the conditions under which they were originally trained. Figure 8 provides representative examples of all three validation approaches, which are explained in more detail below: training set restrictions (Section **??**), cross-dataset generalisation (Section 5.4), and cross-task generalisation (Section 5.5. To assist in

interpretation, the main findings from the generalisation analyses are summarised below before being examined in detail.

Extended versions of these analyses, including reaction time probability density functions (RT-PDFs) and complementary validation cases, are provided in Appendix F (LOCO and cross-dataset analyses) and Appendix G (cross-task generalisation).

5.1 Key findings from generalisation analyses

- LK-based drift diffusion models generalise across experimental conditions, datasets, and task paradigms.
- Core decision and timing structure is robustly preserved under limited training conditions, even when parts of the stimulus space are omitted.
- Generalisation across lab and online datasets is strong, with lab-trained models reproducing online decision timing without refitting, despite the lab data providing sparse coverage of several decision scenarios.
- Distance- and velocity-based DDM formulations transfer reliably to a related task paradigm when refitted, indicating task-level robustness of the evidence accumulation structure.
- Visual augmentation reduces robustness under restricted training and transfer conditions, increasing sensitivity to the specific training data.

5.2 Studied DDMs generalise across conditions, datasets, and task paradigms

The LK-DDMs in the present study were shown to be robust to varying degrees of data quality and generalisable in two distinct senses: in their learned behaviour across datasets of the same task, and in their evidence-accumulation framework across decision tasks. Robustness to sparse and limited data is demonstrated by both the condition-exclusion validation and the cross-dataset evaluation, where models trained on restricted or lower-coverage data retain their ability to reproduce core decision and timing patterns across all conditions. This robustness underscores the explanatory power of the learned behaviour and indicates that LK-DDMs can infer stable decision dynamics from limited training data.

The cross-task validation adds a more fundamental dimension. When the same modelling framework is applied to a different decision task, refitted LK-DDMs achieve comparable predictive performance, despite differences in task structure and experimental context. This result supports the interpretation that the LK-DDMs capture recurrent evidence-accumulation dynamics that are not specific to the narrow passage task, but can be adapted to explain behaviour across multiple related decision scenarios.

5.3 Core decision structure is preserved under limited training conditions

To evaluate robustness under restricted training conditions, leave-one-condition-out (LOCO) analyses were conducted on the online dataset. Two exclusion strategies were considered: removing all trials at a specific initial TTA level, and removing individual TTA–v_{on} combinations. Figure 8.a shows a representative worst-case scenario of a TTA-exclusion trained model used to predict decisions in all conditions. These analyses test whether LK-DDMs require dense coverage of the condition space or whether the decision dynamics can be learned from more parsimonious data and transferred to unseen conditions. Complementary exclusion cases and RT-PDF analyses are reported in Appendix F.

All LOCO-trained models succeeded in reproducing the core patterns in decision outcome and RTs for both lab and online datasets. The predictive precision degraded marginally with all exclusions, and most heavily in the conditions that were excluded during training. Excluding an entire TTA level produced stronger degradation in predictive accuracy than the exclusion of one TTA-v_{on} condition.

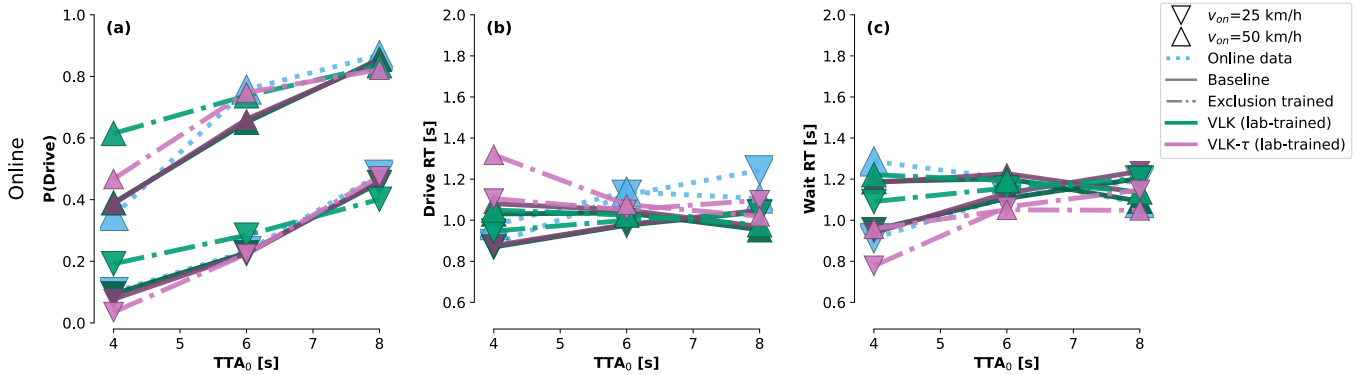
By way of stress-testing the models, the exclusion condition that produced the most divergent results is pictured in Figure 8.a: exclusion of TTA_0=4. The VLK and VLK-τ models are representative because the VLK base model exhibited the strongest overall performance, and its looming-augmented variant allows direct assessment of how visual information affects robustness under restricted training. The decreased predictive accuracy is most apparent in the P(*Drive*) panel, where the VLK model underestimates the sharp transition in P(*Drive*) associated with short TTA conditions. Mean reaction-time predictions are also affected, with reduced curvature across conditions compared to the full-training case. These effects indicate that removing an entire TTA level hinders the models in reproducing the non-linear patterns seen across a range of temporal conditions. Nevertheless, the qualitative relationships between TTA, velocity, decision outcome, and RT remain intact, and even this worst-case exclusion yields plausible decision and timing patterns.

Including visual information generally alters the predictions of unseen conditions across different LOCO-trained models. As illustrated by Figure 8.a, under exclusion training, the looming-augmented model can provide conditional improvement of P(*Drive*) predictions, partially compensating for biases introduced by the missing TTA level. However, this improvement is accompanied by degraded mean RT predictions, especially for *Drive* decisions. Compared to the base model, the VLK-τ model exhibits larger deviations in both *Drive* and *Wait* RTs, indicating increased sensitivity to the TTA-condition coverage of the training data. Thus, visual information slightly stabilises decision outcome predictions of unseen conditions, but reduces the robustness of RT predictions.

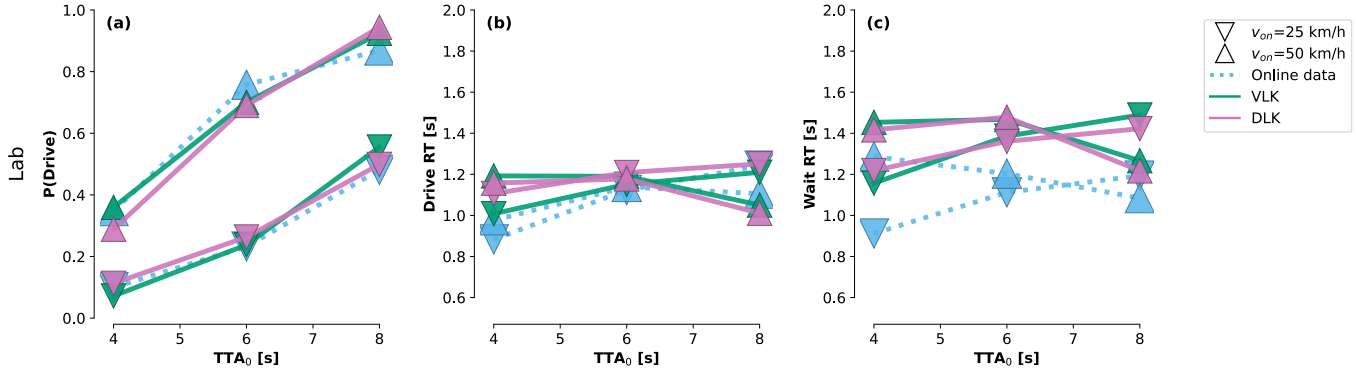
5.4 Cross-dataset generalisation within the narrow passage task

Figure 8b evaluates cross-dataset generalisation within the narrow passage task by evaluating models trained on the lab dataset to the online dataset without refitting. This transfer is non-trivial: the lab data provide substantially fewer observations and sparse coverage of several decision scenarios that are well represented online.

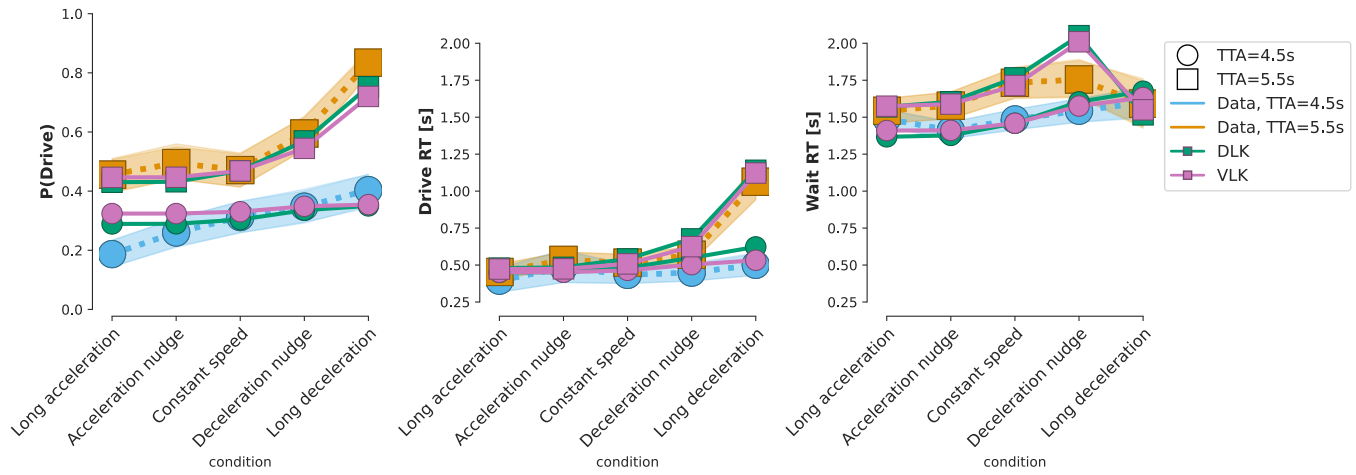
Lab-trained LK-DDMs reproduce the main decision and



(a) Population-averaged *Drive* probability and mean *Drive* and *Wait* RTs from measured online data and predictions of velocity- and velocity-looming-based DDMs trained with all TTA_0=4 conditions excluded. Predictions are shown across the full TTA_0–v_{on} condition grid.



(b) Population-averaged *Drive* probability and mean *Drive* and *Wait* RTs from measured online data and predictions of base models trained on the lab dataset and evaluated on the online dataset without refitting.



(c) Population-averaged *Drive* probability and mean *Drive* and *Wait* RTs from measured left-turn decision data and predictions of base models refitted to the left-turn interaction task. Shaded regions denote 95% confidence intervals of the measured data.

Figure 8: Generalisation and robustness of longitudinal kinematics drift diffusion models (LK-DDMs) across training conditions, datasets, and task paradigms. Each row shows predicted decision proportions (P(Drive)), mean *Drive* RTs, and mean *Wait* RTs (left to right) for a distinct validation scenario. Across all three validations, the models preserve the dominant decision and timing structure despite reduced training data, changes in data quality and coverage, and transfer to a related interaction task. For consistency across validation scenarios and to ensure that excluded conditions remain visually identifiable, all narrow passage plots retain the original TTA_0–v_{on} condition ordering.

timing patterns observed in the online data. Predicted P(*Drive*) from lab-trained models closely follows the online measured relationships with TTA_0 and v_on, and the predicted mean RTs exhibit the same conditional structure as models trained directly on the online dataset. Importantly, this agreement extends to decision scenarios that were weakly represented or absent in the lab training data, indicating that the learned drift structure could

be inferred with limited evidence. Rather than losing predictive accuracy on abundant decisions in favour of irregular data in the more sparse decision scenarios, the models correctly learn that the behaviour governing those data-sparse conditions is synonymous with the more abundant scenarios.

These results demonstrate that the LK-DDM formulations capture decision dynamics that are stable across datasets with markedly different levels of experimental control and behavioural variability. The successful transfer from sparse lab data to richer online data further supports the claim that the fitted models reflect population-level decision structure rather than dataset-specific artefacts.

5.5 Generalisation across task paradigms

Figure 8c presents external validation of the LK-DDMs using data from the left-turn nudging experiment. In this analysis, base LK models were refitted to the left-turn data and evaluated on their ability to reproduce decision proportions and RTs across the task's experimental conditions.

Both distance- and velocity-based LK formulations transfer robustly to this new task context. Predicted P($Drive$) and mean RTs capture the dominant behavioural patterns observed in the Nudging data, with no systematic degradation relative to performance in the narrow passage task. Differences between distance- and velocity- based DDMs remain small and local, mirroring the behaviour observed in the primary task.

This cross-task generalisation indicates that the LK-DDMs are not specific to the narrow passage paradigm, but instead capture a more general evidence-accumulation structure applicable across related interaction contexts.

5.6 Summary of generalisation findings

Across all validation analyses, the LK-based drift diffusion models demonstrate strong robustness and generalisability beyond their original training conditions. Core relationships between kinematics, decision outcome, and reaction timing are preserved when training data are limited, when models are transferred between lab and online datasets, and when drift formulations are refitted to a related task paradigm. Notably, models trained on sparse lab data reproduce decision timing patterns in the larger and more behaviourally variable online dataset without refitting, despite differences in sample size and experimental control.

Together, these results show that the LK-DDM framework captures stable decision dynamics that extend across conditions, datasets, and task contexts.

6 Discussion

6.1 Overview of main findings

This study examined whether longitudinal kinematics-based drift diffusion models (LK-DDMs) can account for driver decision behaviour in a narrow passage task, how the inclusion of visual information in the drift function affected their predictive accuracy, and whether the learned decision dynamics generalise beyond the specific conditions under which the models were trained. The results show that relatively simple LK-based drift formulations capture complex TTA–velocity dependent decision patterns, reproduce key features of reaction-time distributions, and generalise robustly across training exclusions, datasets, and task paradigms. Visual augmentation via looming provides selective improvements to decision proportions, but does not systematically improve reaction-time accuracy or alter the overall accumulation dynamics. Together, these findings demonstrate that parsimonious longitudinal kinematics-based drift diffusion models capture stable and transferable decision dynamics underlying driver behaviour, which generalise across conditions, datasets, and related task contexts.

6.2 Kinematics-dependent drift captures complex interaction decisions

A central finding of this work is that kinematics-dependent drift formulations based on time-to-arrival and distance or velocity provide an effective account of decision behaviour in the narrow passage task. Both distance- and velocity-based LK-DDMs, referred to as DLK and VLK models respectively, reproduce the dominant structure of P($Drive$) and conditional RTs across experimental conditions and across lab and online datasets. Differences between the two formulations are generally modest but systematic. While both model families reproduce the dominant structure of decision and reaction-time behaviour, VLK models show a consistent, if small, advantage in capturing decision patterns, particularly in short-distance conditions.

The small advantage of velocity-based drift formulations over distance-based formulations, together with the selective contribution of incorporating first-order visual information, suggests that drift variants emphasising first-order kinematic signals can capture certain aspects of decision behaviour more accurately. Importantly, these improvements are not global but emerge in specific behavioural regimes, rather than uniformly across conditions. Similar patterns have been reported in other road-interaction contexts, where enriching drift functions with velocity- or acceleration-related information improves the account of particular decision or timing features while preserving the overall qualitative decision structure of the model [16], [18]. At the same time, evidence from related tasks indicates that such higher-order kinematic terms do not universally enhance model performance and may offer limited benefit when evaluated at a global level [17].

Importantly, the present results further show that relatively minimal kinematic structure is sufficient to explain complex interaction decisions in this class of tasks, without requiring additional mechanisms such as collapsing boundaries or explicit urgency signals. This finding is consistent with prior work demonstrating that structured drift dynamics can account for adaptive decision behaviour without the need for additional urgency processes [17], [30], [31].

6.3 TTA-velocity integration is critical in kinematically ambiguous decision regimes

Analysis of reaction time probability density functions (RT-PDFs) indicates that velocity information plays a particularly important role in resolving ambiguity when initial kinematic information is insufficient to support immediate commitment (kinematic ambiguity). Intermediate-distance conditions are characterised by broader RT distributions and increased variability in decision outcome, suggesting that participants delay commitment when the available evidence is less decisive.

In these conditions, longer RTs are systematically associated with shifts in choice probability: higher velocities are associated with an increased likelihood of choosing *Wait*, whereas lower velocities are associated with an increased likelihood of choosing *Drive*. Rather than indicating qualitatively distinct decision modes, this pattern is consistent with prolonged deliberation under perceptual uncertainty, where additional time allows refinement of time-to-arrival and velocity estimates before commitment. Similar associations between increased RTs and uncertainty resolution have been reported in gap-acceptance, pedestrian interaction, and overtaking tasks, where delayed responses are observed under more ambiguous interaction conditions [11], [12], [32].

Differences in RT-PDF shape between *Drive* and *Wait* decisions are consistent with asymmetries in how evidence accumulates toward the two alternatives, but do not support the existence of discrete behavioural regimes. Instead, the observed patterns point to a gradual and continuous accumulation of evidence that is shaped by task demands and uncertainty in kinematic cues [8], [33]. Importantly, longer response times are associated with systematic shifts in choice, indicating that delaying a decision allows additional evidence to be gathered and can influence the eventual outcome.

The present results do not provide enough context to determine the cause of these structural RT-PDF differences. One possibility is that participants deliberately postpone commitment to refine their estimate of the oncoming vehicle's motion. An alternative, non-exclusive explanation is that postponement itself alters the effective evidence available to the decision process, such that the dominant choice becomes increasingly constrained by the evolving kinematic state.

6.4 visual variables provide local but asymmetric contributions to decision proportion and timing

Augmenting the drift function with visual looming yields selective and asymmetric effects. Looming improves P(Drive) predictions in several conditions, particularly at short distances and for *Wait* decisions, indicating that visual expansion cues can refine decision outcome predictions when kinematic information alone is insufficient. However, looming does not consistently improve reaction-time predictions and often sharpens predicted RT distributions without correcting the systematic underestimation of tail mass. These findings align with previous work showing that optical expansion cues are informative in short-distance, high-urgency scenarios [13], [20], [34], but indicate that their contribution is local rather than global.

Moreover, looming-augmented models exhibit reduced robustness under training exclusions, suggesting that visual terms increase sensitivity to the specific composition of the training data.

The apparent discrepancy between strong statistical associations of looming with behaviour and its limited mechanistic impact in the LK-DDMs reflects the difference between correlational and process-level modelling. In accumulation models, kinematic and visual cues compete within a single drift signal, and velocity or distance provides a more stable contribution across the full condition space. Looming, therefore, acts as a contextual modifier of decision commitment rather than a primary driver of accumulation dynamics.

The asymmetric effects of looming on choice and timing suggest that forcing visual expansion cues into the drift term may be suboptimal. Looming improves decision proportions but yields unstable reaction-time predictions, indicating that its influence may act more on decision modulation (i.e., through bias or boundary modulation) than on accumulation rate.

6.5 Generalisability as a core strength of LK-based drift formulations

One of the strongest contributions of this work is the demonstration that LK-based DDMs generalise robustly across training conditions, datasets, and task paradigms. Leave-one-condition-out analyses show that the core decision structure is preserved even when substantial portions of the stimulus space are omitted during training. Cross-dataset evaluation further demonstrates that models trained on sparse lab data reproduce decision timing patterns in the larger and more behaviourally variable online dataset without refitting.

The successful transfer of the LK-DDM framework to the left-turn task provides additional evidence that the inferred drift structure captures task-invariant properties of interaction decision-making rather than task-specific heuristics. Together, these results indicate that the structure of accumulation reflects stable aspects of human decision behaviour in longitudinal interactions. Such robustness strengthens the explanatory value of the LK-DDM framework, demonstrating that its performance does not depend on dense coverage of the stimulus space or task-specific reformulation.

6.6 Population differences and data modality effects

Differences between the lab and online datasets are most apparent in behavioural variability and decisiveness, rather than in the fundamental structure of the decision process. The lab sample exhibits more stable decision tendencies, while the online sample shows greater variability in both RTs and decision proportions. Crucially, however, the way kinematic and visual cues shape decisions remains consistent across populations.

This dissociation between increased variability and preserved decision structure aligns with prior work showing that population characteristics can influence response variability without changing the underlying evidence-accumulation mechanism [35], [36], [37]. Importantly, the strong cross-dataset generalisation observed here supports the use of well-controlled online experiments for behavioural modelling, particularly when large sample sizes are required to capture sparse or ambiguous decision scenarios.

6.7 Limitations and future directions

Several limitations of the present work should be acknowledged, which also point toward promising directions for future research. First, while the LK-based DDMs capture the dominant structure of choice behaviour and mean RTs, they systematically underestimate the tails of reaction-time distributions, particularly *Drive* decisions and in highly ambiguous conditions. This pattern suggests that prolonged deliberation under uncertainty is not fully accounted for by the current drift formulations. Similar limitations have been reported in other perceptual and driving-related accumulation models, and may reflect missing mechanisms related to sustained uncertainty, time-varying evidence reliability, or changes in accumulation noise over time [20], [30], [38]. Addressing these effects may require extensions beyond static drift functions, for example, by allowing evidence quality or noise to evolve as the interaction unfolds.

Second, visual augmentation via looming yields asymmetric improvements, enhancing choice prediction while often degrading reaction-time accuracy. This suggests that forcing visual expansion cues into the drift term may be suboptimal, and that their influence may act through other components of the decision process, such as response bias, boundary setting, or non-decision time. Future work could explicitly compare alternative model formulations in which visual cues modulate different model parameters, thereby testing whether the observed decision–timing trade-offs reflect a genuine partial dissociation or a limitation of the current model order.

Third, all models were fitted at the population level and therefore do not capture stable individual differences in decision bias, response strategy, or sensitivity to kinematic and visual cues. Hierarchical or participant-specific modelling approaches could provide insight into how such differences shape accumulation dynamics and generalisation behaviour, particularly in ambiguous or postponed decision regimes [39], [40]. Incorporating learning effects across trials may further clarify how experience modulates reliance on velocity, distance, or visual information.

Fourth, the experimental condition grid covered a limited set of TTA–velocity combinations, constraining the analysis of non-linear interactions and higher-order kinematic effects. While this design choice enabled controlled generalisation tests, broader sampling of the stimulus space would allow stronger conclusions about how accumulation dynamics scale across more diverse interaction scenarios.

Finally, the absence of direct perceptual or attentional measures limits the interpretation of the cognitive processes underlying delayed decisions. Eye-tracking or gaze-contingent paradigms could help disentangle whether prolonged RTs reflect intentional postponement, perceptual uncertainty, or emergent effects of evolving kinematic constraints. Such measures would be particularly valuable for validating assumptions about how distinct channels of evidence are sampled and updated during longitudinal interactions.

6.8 Implications for driver decision modelling

The present findings demonstrate that relatively simple kinematics-dependent evidence-accumulation models can capture complex interaction behaviour while remaining robust and generalisable. Velocity emerges as a critical cue for resolving interactional ambiguity, whereas visual motion cues such as looming provide context-dependent refinements rather than global improvements. More broadly, the results suggest that direct mappings from perceptual and kinematic variables to drift can provide a parsimonious and scalable basis for modelling driver decision-making.

Notably, the LK-based DDMs reproduce observed behaviour without requiring explicit urgency signals or collapsing boundaries. Instead, moment-to-moment variations in perceptual input and kinematic constraints are sufficient to explain both choice tendencies and reaction-time distributions. This supports modelling approaches in which structure in observed behaviour emerges directly from kinematic and perceptual constraints, rather than being imposed through additional task-independent mechanisms such as explicit urgency signals or collapsing boundaries. In line with previous work on kinematics-dependent decision-making [6], [7], [20], the present results suggest that much of the apparent complexity in driver interaction behaviour can be captured through appropriately structured evidence accumulation without requiring further task-specific or conditional data.

An implication of the present findings for driver decision modelling is that delayed responses in ambiguous interaction scenarios are systematically associated with shifts in decision outcome. The present analyses do not distinguish whether such delays reflect intentional postponement or emergent effects of prolonged evidence accumulation. Importantly, both interpretations are compatible with the LK-based DDM framework: outcome-dependent biases in delayed decisions can arise from continuous evidence accumulation under kinematic uncertainty, without requiring discrete decision modes or additional mechanisms.

From an application perspective, the observed trade-off between choice accuracy and RT accuracy when including visual information in the drift function is not necessarily detrimental. In many driver-interaction and automated-vehicle contexts, accurately predicting decision probabilities is more critical than reproducing precise reaction-time distributions, particularly under high urgency. Notably, in the most urgent conditions the looming-augmented models tend to overestimate mean *Drive* RTs rather than underestimate them, introducing a conservative bias toward later commitment. While not a design objective of the model, such conservativeness may be preferable to optimistic

timing predictions in safety-critical applications, even if it comes at the cost of reduced RT-PDF fidelity.

The present findings also have implications for the role of purely data-driven and machine-learning-based approaches to driver decision modelling. While such models have demonstrated strong predictive performance in many driving scenarios, they often rely on large amounts of task-specific training data and provide limited insight into the mechanisms underlying observed behaviour. The robustness and generalisability of the LK-based DDMs suggest that parsimonious, theory-driven decision models can capture essential behavioural structure with substantially lower data requirements. This positions evidence-accumulation models as a complementary alternative to black-box predictors, particularly in settings where interpretability, extrapolation beyond the training distribution, or integration with cognitive assumptions is required.

More broadly, the results support the use of DDM-based decision agents in traffic and interaction modelling. Because the LK-based DDMs generalise across datasets and task contexts while retaining a compact and interpretable parametrisation, they are well-suited for integration into agent-based simulation frameworks. In this respect, the present findings align with approaches such as the COMMOTIONS framework [41], where driver behaviour is modelled as an evidence-accumulation process embedded within interactive traffic scenarios. The demonstrated robustness of kinematics-dependent drift formulations suggests that such agents can produce realistic and adaptive interaction behaviour without (extensive) retraining or behavioural rules, supporting scalable simulation of heterogeneous traffic environments.

Contributions

This thesis makes the following contributions to the modelling of human driver decision-making in longitudinal interaction scenarios:

- It demonstrates that relatively simple longitudinal kinematics-based drift diffusion models (LK-DDMs) can successfully capture the core choice and timing behaviour observed in narrow-passage interaction decisions, extending prior DDM approaches to a previously unmodelled task.
- It shows that the inferred accumulation dynamics generalise robustly across multiple dimensions, including different sets of experimental conditions, independent datasets, and participant populations, despite substantial differences in behavioural variability and absolute reaction times.
- It quantifies the contribution of adding visual information, including visual angle and looming, to the drift formulation, showing that such cues provide selective and localised improvements-primarily in short-distance *Wait* decisions-rather than global gains in timing accuracy.
- It provides behavioural evidence for delayed driver decision-making in ambiguous conditions, with longer reaction times systematically associated with velocity-dependent shifts in decision proportion, consistent with prolonged evidence accumulation under kinematic uncertainty.

Supplementary materials

Supplementary materials, including data and code for all Narrow Passage analyses, can be found at osf-link.

Acknowledgements

The completion of this work owes thanks in part to the dedicated guidance and support of my supervisors A. Zgonnikov and B. Nallapu. They expertly provided the right snippets of information to stimulate my own inference of fruitful research direction and essential scoping advice that kept me grounded throughout the process. My sincere gratitude goes out to them. This work would not have been possible without the experimental work of M. Baumann and L. Miller, whose input and participation I greatly appreciate. Additionally, I would like to mention the indispensable mental support and discussion partnership from Jesse, Teun, Noah, Jitse, Flint, and Federico, without whom this thesis would have been a much heavier load to carry. Finally, I must thank my parents for their patience and unwavering support in the endeavour that has been my time at the TU Delft. I recognise and cherish my good fortune, which I owe to these and many more people and which enabled me to dedicate so much time and energy to this project.

References

[1] D.-F. Xie, Z.-Z. Fang, B. Jia, and Z. He, "A data-driven lane-changing model based on deep learning," *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 41–60, Sep. 2019, ISSN: 0968-090X. DOI: 10.1016/j.trc.2019.07.002. Accessed: Jan. 4, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X18315651.

[2] C. Wei, F. Hui, and A. J. Khattak, "Driver Lane-Changing Behavior Prediction Based on Deep Learning," en, *Journal of Advanced Transportation*, vol. 2021, no. 1, p. 6 676 092, 2021, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2021/6676092, ISSN: 2042-3195. DOI: 10.1155/2021/6676092. Accessed: Jan. 4, 2026. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2021/6676092.

[3] A. Kuefler, J. Morton, T. Wheeler, and M. Kochenderfer, *Imitating Driver Behavior with Generative Adversarial Networks*, arXiv:1701.06699 [cs], Jan. 2017. DOI: 10.48550/arXiv.1701.06699. Accessed: Jan. 4, 2026. [Online]. Available: http://arxiv.org/abs/1701.06699.

[4] P. G. Gipps, "A behavioural car-following model for computer simulation," *Transportation Research Part B: Methodological*, vol. 15, no. 2, pp. 105–111, Apr. 1981, ISSN: 0191-2615. DOI: 10.1016/0191-2615(81)90037-0. Accessed: Jun. 5, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0191261581900370.

[5] M. Treiber and A. Kesting, "The Intelligent Driver Model with stochasticity – New insights into traffic flow oscillations," *Transportation Research Part B: Methodological*, TRB:ISTTT-22, vol. 117, pp. 613–623, Nov. 2018, ISSN: 0191-2615. DOI: 10.1016/j.trb.2017.08.012. Accessed: Jun. 5, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191261517307014.

[6] G. Markkula et al., "Explaining human interactions on the road by large-scale integration of computational psychological theory," *PNAS Nexus*, vol. 2, no. 6, 2023. DOI: 10.1093/pnasnexus/pgad163.

[7] J. Engström, S.-Y. Liu, A. Dinparastdjadid, and C. Simoiu, "Modeling road user response timing in naturalistic traffic conflicts: A surprise-based framework," *Accident Analysis & Prevention*, vol. 198, p. 107 460, Apr. 2024, ISSN: 0001-4575. DOI: 10.1016/j.aap.2024.107460. Accessed: May 24, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457524000058.

[8] R. Ratcliff and G. McKoon, "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks," *Neural computation*, vol. 20, no. 4, pp. 873–922, Apr. 2008, ISSN: 0899-7667. DOI: 10.1162/neco.2008.12-06-420. Accessed: Apr. 30, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2474742/.

[9] U. Durrani, C. Lee, and D. Shah, "Predicting driver reaction time and deceleration: Comparison of perception-reaction thresholds and evidence accumulation framework," *Accident Analysis & Prevention*, vol. 149, p. 105 889, Jan. 2021, ISSN: 0001-4575. DOI: 10.1016/j.aap.2020.105889. Accessed: May 24, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457520317097.

[10] U. Durrani and C. Lee, "A new car-following model with incorporation of Markkula's framework of sensorimotor control in sustained motion tasks," *Transportation Research Part B: Methodological*, vol. 184, p. 102 969, Jun. 2024, ISSN: 0191-2615. DOI: 10.1016/j.trb.2024.102969. Accessed: May 22, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191261524000936.

[11] A. Zgonnikov, D. A. Abbink, and G. Markkula, "Should I Stay or Should I Go? Cognitive Modeling of Left-Turn Gap Acceptance Decisions in Human Drivers," en, 2022. DOI: 10.1177/00187208221144561. Accessed: May 19, 2025. [Online]. Available: https://repository.tudelft.nl/record/uuid:94f0cff6-5eab-4024-afaf-f078cf03d194.

[12] S. H. A. Mohammad, H. Farah, and A. Zgonnikov, "In the driver's mind: Modeling the dynamics of human overtaking decisions in interactions with oncoming automated vehicles," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 107, pp. 562–577, Nov. 2024, ISSN: 1369-8478. DOI: 10.1016/j.trf.2024.09.020. Accessed: May 21, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1369847824002717.

[13] M. Svärd, G. Markkula, J. Bärgman, and T. Victor, "Computational modeling of driver pre-crash brake response, with and without off-road glances: Parameterization using real-world crashes and near-crashes," *Accident Analysis & Prevention*, vol. 163, p. 106 433, Dec. 2021, ISSN: 0001-4575. DOI: 10.1016/j.aap.2021.106433. Accessed: May 23, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457521004644.

[14] L. Miller, J. Leitner, J. Kraus, and M. Baumann, "Implicit intention communication as a design opportunity for automated vehicles: Understanding drivers' interpretation of vehicle trajectory at narrow passages," *Accident Analysis & Prevention*, vol. 173, p. 106 691, Aug. 2022, ISSN: 0001-4575. DOI: 10.1016/j.aap.2022.106691. Accessed: Dec. 10, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457522001270.

[15] L. Miller, A. Zgonnikov, G. Markkula, S. Jokhio, and M. Baumann, "Modeling Driver Decision-Making in Narrow Passages: An Experimental Study on the Effects of Arrival Time, Speed and Automation Level of Oncoming Vehicles," Ulm University, Germany, 2024.

[16] J. Pekkanen et al., "Variable-Drift Diffusion Models of Pedestrian Road-Crossing Decisions," en, *Computational Brain & Behavior*, vol. 5, no. 1, pp. 60–80, Mar. 2022, ISSN: 2522-087X. DOI: 10.1007/s42113-021-00116-z. Accessed: May 24, 2025. [Online]. Available: https://doi.org/10.1007/s42113-021-00116-z.

[17] A. Zgonnikov, N. Beckers, A. George, D. Abbink, and C. Jonker, "Nudging human drivers via implicit communication by automated vehicles: Empirical evidence and computational cognitive modeling," *International Journal of Human-Computer Studies*, vol. 185, p. 103 224, May 2024, ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2024.103224. Accessed: May 7, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1071581924000089.

[18] M. Theisen, C. Schießl, W. Einhäuser, and G. Markkula, "Pedestrians' road-crossing decisions: Comparing different drift-diffusion models," *International Journal of Human Computer Studies*, vol. 183, 2024. DOI: 10.1016/j.ijhcs.2023.103200.

[19] H. Huang et al., *Understanding Driver Cognition and Decision-Making Behaviors in High-Risk Scenarios: A Drift Diffusion Perspective*, arXiv:2503.12637 [cs], Mar. 2025. DOI: 10.48550/arXiv.2503.12637. Accessed: May 22, 2025. [Online]. Available: http://arxiv.org/abs/2503.12637.

[20] G. Markkula, J. Engström, J. Lodin, J. Bärgman, and T. Victor, "A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies," *Accident Analysis & Prevention*, vol. 95, pp. 209–226, Oct. 2016, ISSN: 0001-4575. DOI: 10.1016/j.aap.2016.07.007. Accessed: May 25, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457516302366.

[21] U. Durrani and C. Lee, "Applying the Accumulator model to predict driver's reaction time based on looming in approaching and braking conditions," *Journal of Safety Research*, vol. 86, pp. 298–310, Sep. 2023, ISSN: 0022-4375. DOI: 10.1016/j.jsr.2023.07.008. Accessed: Nov. 7, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022437523000932.

[22] Renato Contini, "Body Segment Parameters," in *Artificial Limbs: A Review of Current Developments*, ser. 1, vol. 16, Comittee on Prosthetics Research and Development, 1972. Accessed: Dec. 24, 2025. [Online]. Available: http://www.oandplibrary.org/al/pdf/1972_01.pdf.

[23] G. Jocher, J. Qiu, and C. Ayush, *Ultralytics YOLO*, Jan. 2023. [Online]. Available: https://github.com/ultralytics/ultralytics.

[24] H. Schielzeth, "Simple means to improve the interpretability of regression coefficients," en, *Methods in Ecology and Evolution*, vol. 1, no. 2, pp. 103–113,

2010, ISSN: 2041-210X. DOI: 10.1111/j.2041-210X.2010.00012.x. Accessed: Oct. 23, 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2041-210X.2010.00012.x.

[25] A. Agresti, "CHAPTER 5 Models for Binary Data," en, in *Foundations of Linear and Generalized Linear Models*, Newark: John Wiley & Sons, Incorporated, 2015, ISBN: 978-1-118-73005-8. Accessed: Nov. 4, 2025. [Online]. Available: https://learning.oreilly.com/library/view/foundations-of-linear/9781118730058/c05.xhtml.

[26] R. Ratcliff, "A theory of memory retrieval.," en, *Psychological Review*, vol. 85, no. 2, pp. 59–108, Mar. 1978, ISSN: 1939-1471, 0033-295X. DOI: 10.1037/0033-295X.85.2.59. Accessed: Apr. 21, 2025. [Online]. Available: https://doi.apa.org/doi/10.1037/0033-295X.85.2.59.

[27] G. Markkula, "Modeling driver control behavior in both routine and near-accident driving," EN, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, no. 1, pp. 879–883, Sep. 2014, Publisher: SAGE Publications Inc, ISSN: 1071-1813. DOI: 10.1177/1541931214581185. Accessed: May 25, 2025. [Online]. Available: https://doi.org/10.1177/1541931214581185.

[28] M. Svärd, J. Bärgman, and T. Victor, "Detection and response to critical lead vehicle deceleration events with peripheral vision: Glance response times are independent of visual eccentricity," en, *Accident Analysis & Prevention*, vol. 150, p. 105 853, Feb. 2021, ISSN: 00014575. DOI: 10.1016/j.aap.2020.105853. Accessed: Sep. 25, 2025. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0001457520316730.

[29] Maxwell Shinn, Daniel B. Ehrlich, Daeyeol Lee, John D. Murray, and Hyojung Seo, "Confluence of Timing and Reward Biases in Perceptual Decision-Making Dynamics — Journal of Neuroscience," *The Journal of Neuroscience*, pp. 7326–7342, Sep. 2020. DOI: 10.1523. Accessed: Nov. 6, 2025. [Online]. Available: https://doi.org/10.1523/JNEUROSCI.0544-20.2020.

[30] G. E. Hawkins, E.-J. Wagenmakers, R. Ratcliff, and S. D. Brown, "Discriminating evidence accumulation from urgency signals in speeded decision making," *Journal of Neurophysiology*, vol. 114, no. 1, pp. 40–47, Jul. 2015, ISSN: 0022-3077. DOI: 10.1152/jn.00088.2015. Accessed: Apr. 29, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4495756/.

[31] P. L. Smith and R. Ratcliff, "Modeling Evidence Accumulation Decision Processes Using Integral Equations: Urgency-gating and Collapsing Boundaries," *Psychological review*, vol. 129, no. 2, pp. 235–267, Mar. 2022, ISSN: 0033-295X. DOI: 10.1037/rev0000301. Accessed: Apr. 29, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8857294/.

[32] F. Bontje and A. Zgonnikov, "How Sure is the Driver? Modelling Drivers' Confidence in Left-Turn Gap Acceptance Decisions," *Computational Brain and Behavior*, vol. 7, no. 3, pp. 437–456, 2024. DOI: 10.1007/s42113-024-00207-7.

[33] G. E. Hawkins and A. Heathcote, "Racing against the clock: Evidence-based versus time-based decisions," eng, *Psychological Review*, vol. 128, no. 2, pp. 222–263, Mar. 2021, ISSN: 1939-1471. DOI: 10.1037/rev0000259.

[34] Q. Xue, G. Markkula, X. Yan, and N. Merat, "Using perceptual cues for brake response to a lead vehicle: Comparing threshold and accumulator models of visual looming," *Accident Analysis & Prevention*, vol. 118, pp. 114–124, Sep. 2018, ISSN: 0001-4575. DOI: 10.1016/j.aap.2018.06.006. Accessed: May 22, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457518302239.

[35] R. Ratcliff and Y. Vanunu, "The Effect of Aging on Decision Making while Driving: A Diffusion Model Analysis," *Psychology and aging*, vol. 37, no. 4, pp. 441–455, Jun. 2022, ISSN: 0882-7974. DOI: 10.1037/pag0000690. Accessed: May 26, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9677511/.

[36] D. G. Lee and M. Usher, "Value certainty in drift-diffusion models of preferential choice," eng, *Psychological Review*, vol. 130, no. 3, pp. 790–806, Apr. 2023, ISSN: 1939-1471. DOI: 10.1037/rev0000329.

[37] R. Bieri, M. Jäger, N. Gruber, T. Nef, R. M. Müri, and U. P. Mosimann, "A novel computer test to assess driving-relevant cognitive functions–a pilot study," eng, *International Psychogeriatrics*, vol. 26, no. 2, pp. 229–238, Feb. 2014, ISSN: 1741-203X. DOI: 10.1017/S104161021300183X.

[38] S. D. Brown and A. Heathcote, "The simplest complete model of choice response time: Linear ballistic accumulation," *Cognitive Psychology*, vol. 57, no. 3, pp. 153–178, Nov. 2008, ISSN: 0010-0285. DOI: 10.1016/j.cogpsych.2007.12.002. Accessed: May 31, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010028507000722.

[39] R. Ratcliff, "Modeling one-choice and two-choice driving tasks," en, *Attention, Perception, & Psychophysics*, vol. 77, no. 6, pp. 2134–2144, Aug. 2015, ISSN: 1943-393X. DOI: 10.3758/s13414-015-0911-8. Accessed: May 7, 2025. [Online]. Available: https://doi.org/10.3758/s13414-015-0911-8.

[40] C. N. White, R. A. Curl, and J. F. Sloane, "Using Decision Models to Enhance Investigations of Individual Differences in Cognitive Neuroscience," *Frontiers in Psychology*, vol. 7, p. 81, Feb. 2016, ISSN: 1664-1078. DOI: 10.3389/fpsyg.2016.00081. Accessed: May 6, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4746304/.

[41] A. R. Srinivasan et al., *The COMMOTIONS Urban Interactions Driving Simulator Study Dataset*, arXiv:2305.11909 [cs], Jul. 2024. DOI: 10.48550/arXiv.2305.11909. Accessed: Jun. 3, 2025. [Online]. Available: http://arxiv.org/abs/2305.11909.

A Visual angle reconstruction and data transformation

This appendix describes the geometric assumptions and signal-processing steps used to reconstruct visual angle and looming-related variables from the recorded video stimuli. The purpose is to document the transformation from image-space measurements to visual predictors used in the decision models.

A.1 Viewing geometry and visual-angle approximation

Figure 9 illustrates the experimental viewing geometry and the construction lines used to approximate perspective geometries. Participants were seated at a known fixed horizontal distance from the display, which was assumed to be equal to the horizontal distance between the eyes and the display d_{eye}. The participant's eye position was approximated as a single fixed point between the eyes. The eye height was determined with respect to the tabletop according to German anatomical measurements from Renato Contini [22].

From the construction lines in Figure 9 and orthogonality assumptions, the relative distances were approximated between the display centre and the eye point. The video frame-to-display ratio was known, and the dimensions of the display were known, allowing for the construction of a mapping from video pixel coordinates to spatial coordinates on the display, to rays originating from the eye point. These, in turn, were used to calculate visual angles.

This derivation assumes no head movement throughout different trials, and approximation was used solely for reconstructing visual variables from recorded video. Therefore, the calculations did not affect other measures from the raw data.

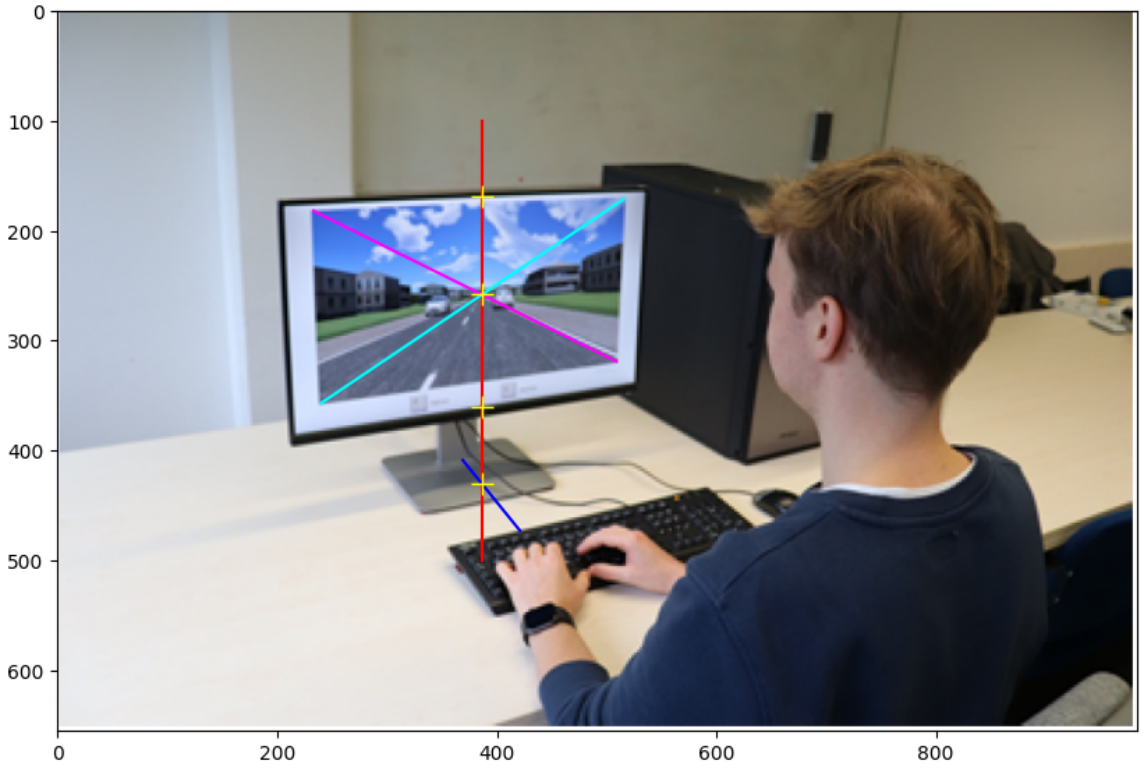


Figure 9: Experimental viewing geometry and construction lines used for visual-angle approximation. Distances between the participant's eyes, the display, and image features were used to map image measurements to visual angles.

A.2 Object detection and bounding-box extraction

The oncoming vehicle was detected in each video frame using an object-detection pipeline, yielding a rectangular bounding box enclosing the vehicle with the use of YOLOv8, a publicly available object detection algorithm. This version of YOLO was chosen because its performance was highly similar to state-of-the-art versions, but the processing time remains relatively small [23]. Figure 10 shows a representative frame with the detected bounding box.

A.3 Temporal fitting of bounding-box dimensions

Raw bounding-box dimensions exhibited frame-to-frame noise and missing values due to detection uncertainty and imperfections. To obtain smooth time-series, the bounding-box dimensions and coordinates were averaged across videos that shared the same kinematic condition and fitted with exponential functions of the form

$$w(t) = a \exp(bt) + c, \tag{7}$$

Figure 10: Representative video frame with detected bounding box around the oncoming vehicle. Bounding-box dimensions provide the image coordinate input for visual angle reconstruction.

where a, b, and c are condition-specific fit parameters. This exponential formulation ensures a positive increasing shape that suits the bounding box height, width, and y-coordinate. The x-coordinate, however, was strictly decreasing, thus the function was fitted to $-x$ values. Other types of functions were fitted to determine which performed best, and the exponential formulation performed best.

Figure 11 shows representative averaged bounding-box trajectories together with the corresponding exponential fits. These fits allowed for the computation of smooth time derivatives, used for the calculation of visual looming.

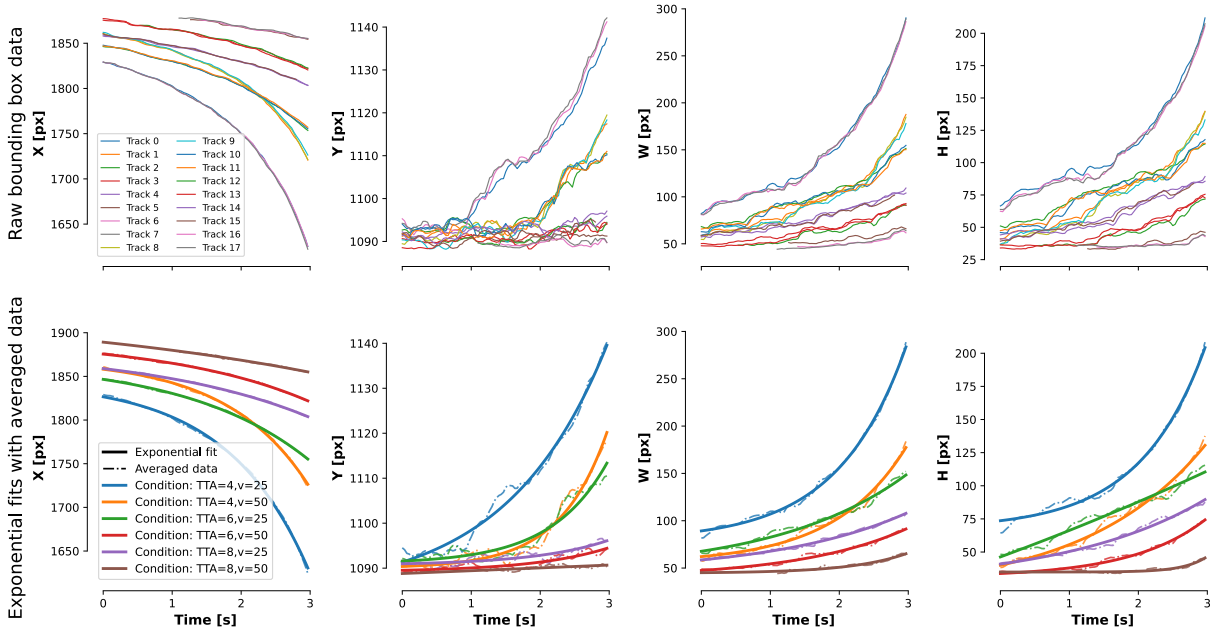


Figure 11: **Top:** Tracked bounding box dimensions and coordinates for all tracked videos. **Left:** Condition-averaged bounding-box dimensions (dotted) and fitted exponential functions (solid). Fitted exponentials were used to generate smooth and complete time series for bounding box dimensions.

A.4 Visual angle and looming reconstruction

To compute the horizontal visual angle subtended by the oncoming vehicle, the object's bounding-box width was reconstructed in three-dimensional space and projected into angular coordinates relative to the participant's eye position. Let (x_l, x_r) denote the left and right edges of the bounding box, (y_t, y_b) its top and bottom edges, and $z(t)$ the depth of the object at time t. The eye position is given by $\mathbf{e} = (e_x, e_y, 0)$. The vertical midpoint of the bounding

box is defined as $y_c = \frac{1}{2}(y_t + y_b)$. Two ray vectors originating from the eye and intersecting the left and right midpoints of the bounding box are then constructed as

$$\mathbf{r}_{\text{left}}(t) = \begin{pmatrix} x_l - e_x \\ y_c - e_y \\ z(t) \end{pmatrix}, \qquad \mathbf{r}_{\text{right}}(t) = \begin{pmatrix} x_r - e_x \\ y_c - e_y \\ z(t) \end{pmatrix}.$$

The instantaneous horizontal visual angle $\theta_w(t)$ is defined as the angle between these two rays and is computed using the cosine formula

$$\theta_w(t) = \arccos\left(\frac{\mathbf{r}_{\text{left}}(t) \cdot \mathbf{r}_{\text{right}}(t)}{\|\mathbf{r}_{\text{left}}(t)\| \, \|\mathbf{r}_{\text{right}}(t)\|} \right).$$

This formulation yields the object's angular width as perceived from the eye, without relying on small-angle or pinhole-camera approximations, and directly reflects the geometric structure of the rendered video stimuli.

Visual looming was quantified as the instantaneous rate of relative expansion of the object's horizontal visual angle. Let $\theta_w(t)$ denote the horizontal visual angle defined above as a function of time. Visual looming was computed using:

$$\tau^{-1}(t) = \frac{\dot{\theta}_w(t)}{\theta_w(t)},$$

where $\dot{\theta}_w(t)$ denotes the temporal derivative of the visual angle. This quantity captures the proportional rate of change of the object's angular size and provides a scale-invariant measure of visual expansion that increases nonlinearly as the object approaches the observer.

A.5 Implementation of visual information in the model fitting pipeline

For implementation of visual variables in the model fitting pipeline, both $\theta(t)$ and $\tau^{-1}(t)$ had to be formulated by a single function of time. Therefore, they were summarised using logistic functions fitted to the computed time-series,

$$f(t) = \frac{L}{1 + \exp[-k(t - t_0)]}, \tag{8}$$

where L controls the amplitude, k the slope, and t_0 the inflection point. These fits provide smooth and compact representations of the temporal evolution of visual variables while preserving their characteristic growth profiles. Other types of functions were fitted to determine which performed best, and the logistic formulation performed best.

Figure 12 shows the manually computed visual-angle and looming values, together with the fitted logistic functions. All transformations operate deterministically on the recorded video data and do not introduce additional behavioural assumptions beyond those inherent in the visual reconstruction procedure.

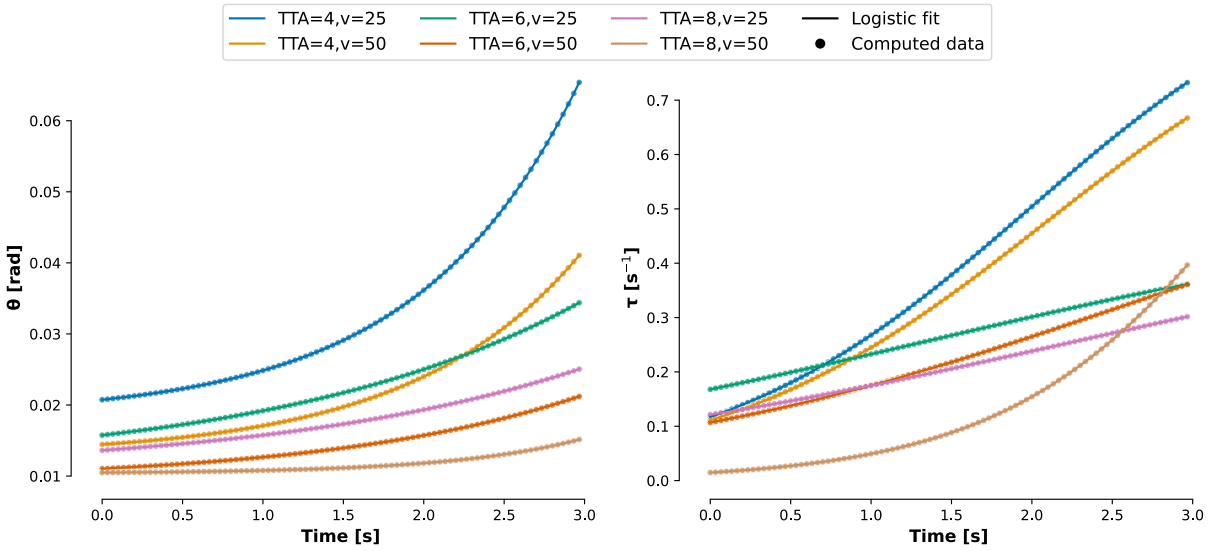


Figure 12: Visual angle and looming over time, derived from fitted bounding-box dimensions and from fitted logistic functions. Points indicate values computed from averaged image-space measurements, while lines show logistic fits used to obtain smooth temporal representations for modelling.

B Reference experiment and modelling summary

This appendix briefly summarises the experimental paradigm, behavioural analyses, and decision-model formulations reported in Zgonnikov et al. [17], which serves as the reference study for the present thesis, and validation environment of the generalisability of the model frameworks in Section 5. Only key aspects relevant for contextualising the modelling transfer are reported.

B.1 Experimental paradigm

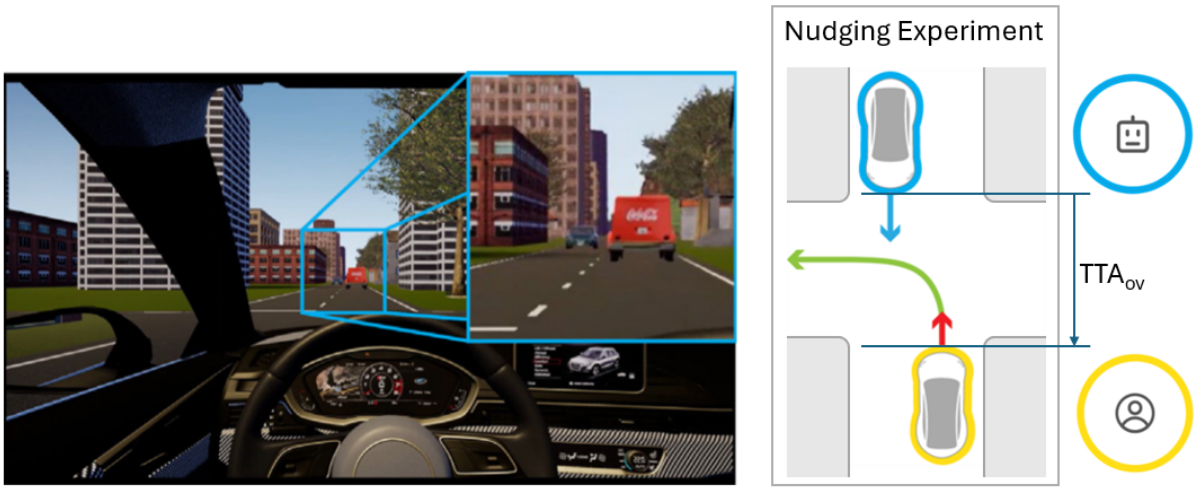


Figure 13: POV (left) and schematic overview (right) of the unprotected left-turn interaction paradigm used in the reference study.

The reference study examined driver decision-making in an *unprotected left-turn* scenario involving an oncoming automated vehicle (AV). Participants drove an ego vehicle and repeatedly decided whether to turn *before* (*Go*) or *after* (*Stay*) the AV passed the intersection. The AV always had right of way but executed subtle longitudinal manoeuvres (deceleration or acceleration nudges) designed to implicitly influence the driver's decision without explicitly yielding.

Initial time-to-arrival (TTA_0) and AV manoeuvre condition were systematically varied, creating interaction scenarios with differing urgency and ambiguity. Decision outcome and reaction time were recorded, along with occasional subjective ratings of AV behaviour.

B.2 Statistical analysis of behaviour

Table 9: Coefficients of the mixed-effects logistic regression describing the final decision as a function of TTA_0 and condition (with Constant speed as the reference category). Participant ID was included as a random intercept. From [17]

	b	SE	z	p
(Intercept)	-9.30	0.71	-13.10	< 0.001
TTA_0	1.82	0.10	17.74	< 0.001
Long acceleration	-0.78	0.15	-5.38	< 0.001
Acceleration nudge	-0.13	0.14	-0.94	0.35
Deceleration nudge	0.50	0.14	3.44	< 0.001
Long deceleration	1.50	0.15	9.89	< 0.001

Decision outcomes were analysed using mixed-effects logistic regression with TTA_0 and AV manoeuvre condition as fixed effects and participant ID as a random intercept. Reaction times were analysed using linear mixed-effects models, complemented by Type-III ANOVA to summarise the contribution of experimental factors and their interactions.

The analyses showed that both choice and timing were systematically modulated by TTA_0 and the AV's longitudinal behaviour, with deceleration nudges increasing the likelihood of *Go* decisions. Reaction times exhibited interaction-dependent delays, indicating prolonged decision processes under specific combinations of kinematic conditions and manoeuvres.

B.3 Decision model formulations

Behaviour was modelled using kinematics-dependent drift diffusion models in which evidence accumulation was driven primarily by time-to-arrival and distance-related variables. Model variants differed in boundary formulation, inclusion of an initial bias, and the presence of an explicit acceleration term in the drift rate.

Table 10: ANOVA table based on the mixed-effects linear regression describing response time as a function of decision, TTA_0, and condition. From [17].

	SS	MS	df	F	p
Decision	5.27	5.27	1	19.54	< 0.001
TTA_0	11.79	11.79	1	43.73	< 0.001
Condition	3.25	0.81	4	3.01	0.02
Decision: TTA_0	0.03	0.03	1	0.10	0.75
Decision: condition	7.71	1.93	4	7.15	< 0.001
Condition: TTA_0	4.55	1.14	4	4.22	0.002
Decision:condition: TTA_0	8.41	2.10	4	7.80	< 0.001

Table 11: Model comparison results reported in the reference study. Eight decision-model configurations were evaluated, differing in boundary type (constant vs. collapsing), presence of an initial bias, and inclusion of an acceleration term. Models are compared using the Bayesian Information Criterion (BIC), where lower values indicate better fit after accounting for model complexity. Model 2 (constant boundary, initial bias, no acceleration term) achieved the lowest BIC and was selected as the best-performing model.

	Without acceleration term		**With acceleration term**	
	No initial bias	With initial bias	No initial bias	With initial bias
Constant boundary				
Model 1	$\kappa = 6$ BIC = 5050	$\kappa = 7$ **BIC = 4610**	$\kappa = 7$ BIC = 5063	$\kappa = 8$ BIC = 4620
Collapsing boundary				
Model 3	$\kappa = 8$ BIC = 5072	$\kappa = 9$ BIC = 4622	$\kappa = 9$ BIC = 5090	$\kappa = 10$ BIC = 4638

Model comparison based on BIC favoured a parsimonious formulation with a constant decision boundary and an initial bias, without an explicit acceleration term. Despite the use of acceleration-based nudges, behavioural effects were explained through their impact on evolving TTA rather than through separate acceleration-sensitive accumulation dynamics.

B.4 Relevance for the present thesis

The left-turn study provides independent evidence that interaction-level driver decisions can be captured by continuous evidence accumulation driven by kinematic variables. Its findings support the use of parsimonious, kinematics-dependent drift formulations and motivate the cross-task generalisation of longitudinal kinematics-based DDMs evaluated in Section 5.5.

C Extended statistical analysis

This appendix provides extended statistical results that complement the analyses reported in Sections 2.3.2 and 3. The material included here serves to document full model specifications, coefficient estimates, and model-comparison tests that are summarised in the main text. No additional conclusions are drawn from these results beyond those already reported. All relevant R code can be found at osf-link.

C.1 Full reaction-time model coefficients

Table 12 reports the complete set of fixed-effect coefficient estimates for the reaction-time mixed-effects model discussed in Section 2.3. The table includes all main effects and interaction terms retained in the final model specification.

These coefficients are provided for completeness and reproducibility. Interpretation of the dominant effects and interaction patterns is given in the main text; the extended table allows inspection of effect magnitudes and uncertainty estimates for all predictors.

Table 12: Fixed effects from the linear mixed-effects models predicting reaction time in the lab and online datasets (REML).

	Lab dataset					**Online dataset**				
Term	Est.	SE	df	t	p	Est.	SE	df	t	p
(Intercept)	1.302	0.017	36.8	76.84	***	1.216	0.008	176.7	148.75	***
decision.e	-0.050	0.004	8472	-13.99	***	-0.013	0.001	40140	-9.114	***
TTA_0.c	0.314	0.002	8471	140.93	***	0.288	0.001	40110	281.37	***
$\tau^{-1}.z$	0.776	0.004	8358	180.70	***	0.769	0.002	39480	350.57	***
v_{on}	0.262	0.003	8468	77.17	***	0.241	0.001	40110	173.34	***
decision.e : TTA_0.c	0.035	0.002	8457	14.48	***	0.029	0.001	40060	29.35	***
decision.e : $\tau^{-1}.z$	0.061	0.004	8454	16.10	***	0.034	0.002	40050	20.01	***
decision.e : v_{on}	-0.011	0.004	8453	-3.099	**	-0.003	0.001	40050	-1.905	.
TTA_0.c : $\tau^{-1}.z$	0.108	0.002	8461	56.35	***	0.098	0.001	40050	108.94	***
$\tau^{-1}.z$: v_{on}	-0.092	0.003	8450	-27.22	***	-0.076	0.002	40030	-47.19	***
TTA_0.c : v_{on}	0.056	0.002	8456	22.61	***	0.074	0.001	40050	73.97	***

Significance codes: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$.

Type-III ANOVA results

To assess the relative contribution of predictors and interaction terms in the reaction-time model, a Type-III ANOVA was conducted. Table 13 summarises the resulting test statistics.

The ANOVA confirms that temporal and kinematic predictors account for the majority of explained variance in reaction time, with several higher-order interactions contributing significantly. These results support the reduced interpretation presented in Section 2.3.

Table 13: Type-III ANOVA tables for the final mixed-effects models describing reaction time in the lab and online datasets (REML).

	Lab dataset						**Online dataset**					
Term	SS	MS	NumDF	DenDF	F	p	SS	MS	NumDF	DenDF	F	p
decision.e	8.47	8.47	1	8471.60	195.80	***	3.75	3.75	1	40141.43	83.06	***
TTA_0.c	859.19	859.19	1	8470.76	19860.75	***	3573.62	3573.62	1	40114.76	79169.51	***
$\tau^{-1}.z$	1412.48	1412.48	1	8358.22	32650.60	***	5547.49	5547.49	1	39477.67	122898.22	***
v_{on}	257.62	257.62	1	8468.28	5955.08	***	1356.33	1356.33	1	40107.85	30047.95	***
decision.e : TTA_0.c	9.07	9.07	1	8456.60	209.60	***	38.88	38.88	1	40060.70	861.42	***
decision.e : $\tau^{-1}.z$	11.21	11.21	1	8454.41	259.06	***	18.08	18.08	1	40051.11	400.57	***
decision.e : v_{on}	0.42	0.42	1	8452.80	9.61	**	0.16	0.16	1	40050.49	3.63	.
TTA_0.c : $\tau^{-1}.z$	137.37	137.37	1	8460.65	3175.30	***	535.70	535.70	1	40046.93	11867.72	***
$\tau^{-1}.z$: v_{on}	32.05	32.05	1	8449.52	740.78	***	100.50	100.50	1	40025.12	2226.48	***
TTA_0.c : v_{on}	22.11	22.11	1	8455.55	511.04	***	246.97	246.97	1	40045.99	5471.23	***

Significance codes: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$.

C.3 Random versus fixed intercepts

To evaluate the necessity of participant-level random intercepts, mixed-effects models with fixed intercepts were compared against models including random intercepts. Model comparison was performed using information criteria, with the intraclass correlation coefficient (ICC) reported as a descriptive measure of between-participant variance. The ICC describes the proportion of variance in the dataset that can be attributed to stable differences between participants.

Table 14 reports the comparison results for both decision and reaction-time models. The inclusion of participant-level random intercepts captures substantial between-participant variability, consistent with the modelling assumptions adopted in Section 2.3.

Table 14: Significance of random intercept inclusion and intra-class correlations (ICC) for lab and online models.

Dataset	**Model type**	$\chi^2(1)$	**p**	**ICC**
Lab	Reaction time (LMM)	998.05	***	0.18
Lab	Decision outcome (GLMM)	905.58	***	0.25
Online	Reaction time (LMM)	5045.09	***	0.20
Online	Decision outcome (GLMM)	7804.08.22	***	0.37

Significance codes: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

C.2 Random versus fixed slopes

An additional set of model comparisons tested whether allowing random slopes for key predictors improved model fit relative to fixed-slope formulations. These comparisons were conducted for the reaction-time model only, as continuous RT measures provide greater information for estimating participant-specific effects and allow more reliable identification of random-slope variance. Equivalent extensions of the binary decision model led to singular or non-convergent fits and were therefore not pursued further.

Table 15 summarises the results of these comparisons. Allowing random slopes for selected predictors yielded only marginal improvements in model fit and did not materially alter population-level fixed effects. Accordingly, fixed-slope specifications were retained to maintain model parsimony and interpretability.

Table 15: Likelihood ratio test comparing random-intercept and random-slope models for reaction time (lab dataset).

	n_{par}	AIC	BIC	logLik	$-2 * \log(L)$	χ^2	Df	p
Lab								
$model_{fixed}$	13.00	-26.24	65.36	26.12	-52.24			
$model_{randslopes}$	22.00	-812.78	-657.76	428.39	-856.78	804.54	9	< 0.001
Online								
$model_{fixed}$	13.00	-2672.95	-2561.14	1349.47	-2698.95			
$model_{randslopes}$	22.00	-7802.68	-7613.47	3923.34	-7846.68	5147.73	9	< 0.001

D Extended model definitions and full-data results

This appendix provides extended documentation and diagnostic results for the drift diffusion models evaluated in the present thesis. It complements the model comparisons and summary figures reported in the main text by listing all tested model definitions, full-data parameter estimates, and additional visualisations of predictions of the focus models in the present study, under full-data training. The focus models include six longitudinal kinematics drift diffusion models (LK-DDMs) separated into two classes: distance-based LK-DDMs (DLK models), and velocity-based LK-DDMS (VLK models). They are optionalyy extended by adding visual angle (θ) or visual looming (τ). The material is included for completeness and transparency and does not introduce new modelling claims.

D.1 Model definitions

Table 16 lists all tested decision-model configurations, including variations in drift formulation, boundary structure, and sensory extensions. Models are identified by their numerical labels, which are used consistently throughout the main text, appendices, and file naming conventions.

This table serves as a reference for interpreting subsequent parameter estimates and model-output figures.

Table 16: Overview of all drift–diffusion model variants used in the modelling analysis. Models 4, 0, 10, 18, 9, and 7 correspond to the six focus models (labelled Focus 1–6).

Model	**Drift** $s(t)$	**Boundary**	**Bias** x_0	**Non-decision time**	**Predictor order**	
					Kinematic	Visual
0 (VLK)	$\alpha(\mathrm{TTA} + \beta_v v - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1	-
1	$\alpha(\mathrm{TTA} + \beta_v v - \theta_{crit})$	B	0	$(\mu_{NDT}, \sigma_{NDT})$	1	-
2	$\alpha(\mathrm{TTA} + \beta_v v - \theta_{crit})$	$\frac{b_0}{1+e^{-k(\mathrm{TTA}(t)-TTA_{\mathrm{crit}})}}$	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1	-
3	$\alpha(\tau^{-1} + \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	-	1
4 (DLK)	$\alpha(\mathrm{TTA} + \beta_d d - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	-
5	$\alpha(\mathrm{TTC} + \beta_v v - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1	-
6	$\alpha(\mathrm{TTA} + \beta_\tau \tau^{-1} - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	1
7 (VLK-τ)	$\alpha(\mathrm{TTA} + \beta_\tau \tau^{-1} + \beta_v v - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1	1
8	$\alpha(\mathrm{TTA} + \beta_v v + \beta_d d - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1, 0	-
9 (DLK-τ)	$\alpha(\mathrm{TTA} + \beta_\tau \tau^{-1} + \beta_d d - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	1
10 (DLK-θ)	$\alpha(\mathrm{TTA} + \beta_\theta \theta_v + \beta_d d - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	0
11	$\alpha(\mathrm{TTA} + \beta_d d - \theta_{crit}) + \beta_\theta \theta_v$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	0
12	$\alpha(\mathrm{TTA} + \beta_d d - \theta_{crit}) - \beta_\theta \theta_v$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	0
13	$\alpha(\beta_v v + \beta_d d - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1, 0	-
14	$\alpha(\beta_v v + \beta_\theta \theta_v - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1	0
15	$\alpha(\beta_\tau \tau^{-1} + \beta_\theta \theta_v - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	-	1, 0
16	$\beta_v v - \theta$	B	$x_0 + s\,d$	$(\mu_{NDT}, \sigma_{NDT})$	1	-
17	$\beta_v v - \theta$	B	$x_0 + s\,\theta_v$	$(\mu_{NDT}, \sigma_{NDT})$	1	0
18 (VLK-θ)	$\alpha(\mathrm{TTA} + \beta_\theta \theta_v + \beta_v v - \theta_{crit})$	B	x_0	$(\mu_{NDT}, \sigma_{NDT})$	1	0
19	$\alpha(\mathrm{TTA} + \beta_d d - \theta_{crit})$	B	0	$(\mu_{NDT}, \sigma_{NDT})$	0	-
20	$\alpha(\mathrm{TTA} + \beta_d d - \theta_{crit})$	$\frac{b_0}{1+e^{-k(\mathrm{TTA}(t)-TTA_{\mathrm{crit}})}}$	x_0	$(\mu_{NDT}, \sigma_{NDT})$	0	-

D.2 Full-data parameter estimates

Tables 17 and 18 report parameter estimates for all models trained on the full laboratory and online datasets, respectively. These estimates correspond to the models evaluated in the main text and are provided here to allow inspection of fitted parameter values across datasets.

Parameter estimates are shown separately for laboratory and online training to reflect differences in data volume and experimental context. Interpretation of parameter trends and model performance is given in the main text; the tables included here serve as detailed reference material.

D.3 Distance-ordered predictions for online-trained VLK models

Figure 14 presents extended diagnostic visualisations for the VLK model family (Models 0, 18, and 7) trained on the full online dataset. The figure combines aggregate behavioural measures with distance-ordered reaction-time distributions, providing a complementary view to the summary results shown in the main text.

The top panel reports condition-level predictions and empirical values for the probability of choosing to Drive, mean Drive reaction time, and mean Wait reaction time. The bottom panel shows distance-ordered reaction-time PDFs for

Table 17: BIC values and fitted parameters for all drift–diffusion models trained on the lab dataset. Only parameters used by a model are shown. β_{TTA} is omitted because it is fixed at 1 for all models.

Model	Focus	BIC	Fitted parameters
0	–	23045.49	α=0.3628, β_v=0.4114, θ=9.5704, B=1.0872, x_0=0.0059, μ=0.3679, σ=0.0916
1	–	23038.38	α=0.3646, β_v=0.4108, θ=9.5470, B=1.1040, μ=0.3591, σ=0.0879
2	–	22798.55	α=0.3254, β_v=0.4586, θ=9.8306, x_0=−0.0323, μ=0.3378, σ=0.0744, b_0=1.1742, k=1.8883, TTA_{crit}=2.0320
3	–	25146.11	α=7.4823, θ=0.1890, B=1.0486, x_0=−0.1308, μ=0.3569, σ=0.0937
4	–	23186.52	α=0.0872, β_d=0.2941, θ=20.9569, B=1.0710, x_0=−0.0164, μ=0.3765, σ=0.0962
5	–	23174.73	α=1.2039, β_v=0.0953, θ=5.3030, B=1.0889, x_0=−0.0382, μ=0.3666, σ=0.0927
6	–	25162.87	α=0.3004, β_τ=0, θ=5.4088, B=1.0228, x_0=0.0420, μ=0.3840, σ=0.0936
7	–	23000.14	α=0.4328, β_v=0.3942, β_τ=5.2595, θ=10.4187, B=1.0915, x_0=0.0600, μ=0.3744, σ=0.0945
9	–	23002.94	α=0.1332, β_τ=32.4511, β_d=0.2549, θ=25.1914, B=1.0846, x_0=0.0811, μ=0.3895, σ=0.1027
10	–	23201.60	α=0.0865, β_d=0.2978, β_{ttc}=9.7701, θ=21.3126, B=1.0658, x_0=−0.0148, μ=0.3796, σ=0.0976
11	–	23208.06	α=0.0822, β_d=0.3115, β_{ttc}=0, θ=22.5415, B=1.0823, x_0=0.0362, μ=0.3733, σ=0.0937
13	–	23925.98	α=0.0354, β_d=0.6989, θ=50, B=0.9258, x_0=0.0673, μ=0.5032, σ=0.1778
14	–	25718.79	α=0.1189, β_θ=0, θ=10.7354, B=1.0060, x_0=0.0729, μ=0.3880, σ=0.0933
15	–	27105.31	α=0.0015, β_θ=0, θ=44.0253, B=0.9976, x_0=0.0958, μ=0.3843, σ=0.0896
16	–	28411.33	β_v=2.6414, θ=10.8601, B=1.0233, μ=1.2889, σ=0.5, x_0^{lin}=−0.7747, s=−0.3420
17	–	25727.58	β_v=0.1195, θ=1.2661, B=1.0297, μ=0.3738, σ=0.0871, x_0^{lin}=−0.0326, s=0.0556
18	–	23055.84	α=0.3621, β_v=0.4127, β_{ttc}=0.5760, θ=9.6210, B=1.0863, x_0=0.0147, μ=0.3685, σ=0.0917
19	–	23175.10	α=0.0756, β_d=0.3474, θ=24.0478, B=1.0981, μ=0.3626, σ=0.0894
20	–	22971.73	α=0.0556, β_d=0.4890, θ=30.6749, x_0=0.0134, μ=0.3414, σ=0.0931, b_0=1.9994, k=0.1313, TTA_{crit}=3.3639

Table 18: BIC values and fitted parameters for all drift–diffusion models trained on the online dataset. Only parameters used by a model are shown. β_{TTA} is omitted because it is fixed at 1 for all models.

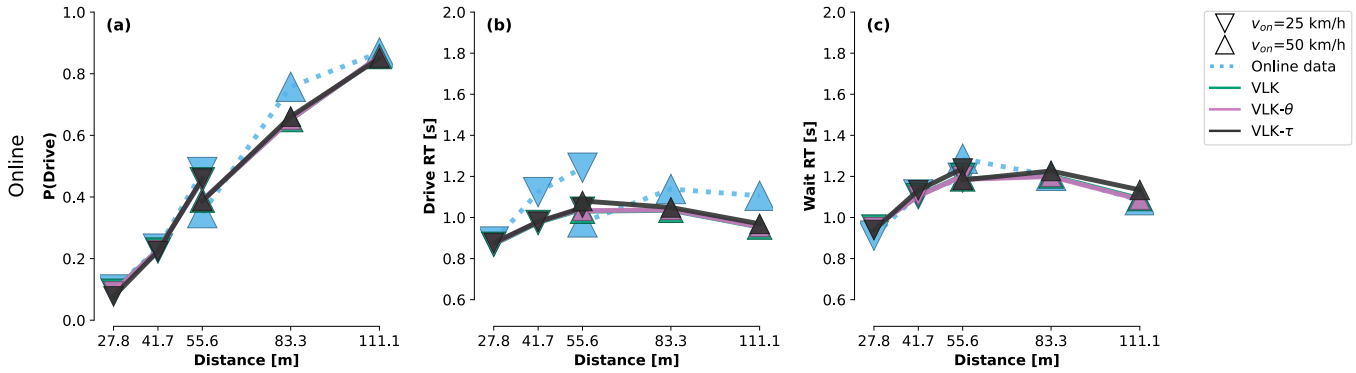
Model	Focus	BIC	Fitted parameters
0	–	101624.74	α=0.2956, β_v=0.5022, θ=11.0274, B=0.9488, x_0=−0.0003, μ=0.3018, σ=0.1124
1	–	101617.86	α=0.2951, β_v=0.5035, θ=11.0610, B=0.9508, μ=0.2986, σ=0.1112
2	–	101900.92	α=0.2833, β_v=0.5935, θ=11.7100, x_0=−0.0399, μ=0.2779, σ=0.1213, b_0=2.0315, k=0.0215, TTA_{crit}=8.0487
3	–	108890.04	α=6.6272, θ=0.1665, B=0.9505, x_0=−0.1256, μ=0.2796, σ=0.1050
4	–	102269.44	α=0.0330, β_d=0.7449, θ=49.6938, B=0.9515, x_0=−0.0196, μ=0.2988, σ=0.1120
5	–	102178.88	α=0.9588, β_v=0.1232, θ=5.7611, B=0.9147, x_0=−0.0297, μ=0.3257, σ=0.1222
6	–	109945.68	α=0.2482, β_τ=0, θ=5.9077, B=0.8969, x_0=0.0254, μ=0.3165, σ=0.1158
7	–	101644.94	α=0.3338, β_v=0.4807, β_τ=3.7269, θ=11.4613, B=0.9782, x_0=0.0289, μ=0.2853, σ=0.1066
9	–	101926.81	α=0.0549, β_τ=42.5910, β_d=0.5311, θ=44.1953, B=0.9579, x_0=0.0143, μ=0.3004, σ=0.1126
10	–	102326.95	α=0.0404, β_d=0.6013, β_{ttc}=2.5547, θ=41.3532, B=0.9695, x_0=−0.0128, μ=0.2891, σ=0.1076
11	–	102279.14	α=0.0326, β_d=0.7564, β_{ttc}=0, θ=50, B=0.9504, x_0=−0.0297, μ=0.2990, σ=0.1115
13	–	102415.32	α=0.0338, β_d=0.6631, θ=50, B=0.9403, x_0=−0.0159, μ=0.3043, σ=0.1132
14	–	108962.71	α=0.1294, β_θ=0, θ=11.4579, B=0.9216, x_0=0.0306, μ=0.3013, σ=0.1095
15	–	115301.78	α=0.0027, β_θ=0, θ=50, B=0.8878, x_0=0.0503, μ=0.3138, σ=0.1131
16	–	128512.57	β_v=2.6209, θ=44.9578, B=1.0216, μ=1.0400, σ=0.5, x_0^{lin}=−0.3216, s=0.7510
17	–	108993.37	β_v=0.1294, θ=1.4697, B=0.9298, μ=0.2978, σ=0.1080, x_0^{lin}=0.0175, s=−0.1750
18	–	101629.82	α=0.2966, β_v=0.5009, β_{ttc}=0.5542, θ=11.0005, B=0.9458, x_0=−0.0069, μ=0.3019, σ=0.1124
19	–	102316.53	α=0.0331, β_d=0.7370, θ=50, B=0.9359, μ=0.3084, σ=0.1152
20	–	102132.37	α=0.0647, β_d=0.3565, θ=26.5391, x_0=−0.0171, μ=0.2949, σ=0.1119, b_0=1.8961, k=0.0520, TTA_{crit}=4.8714

Drive and Wait decisions. Together, these visualisations illustrate how the VLK models reproduce both mean-level trends and distributional timing structure under full-data training.

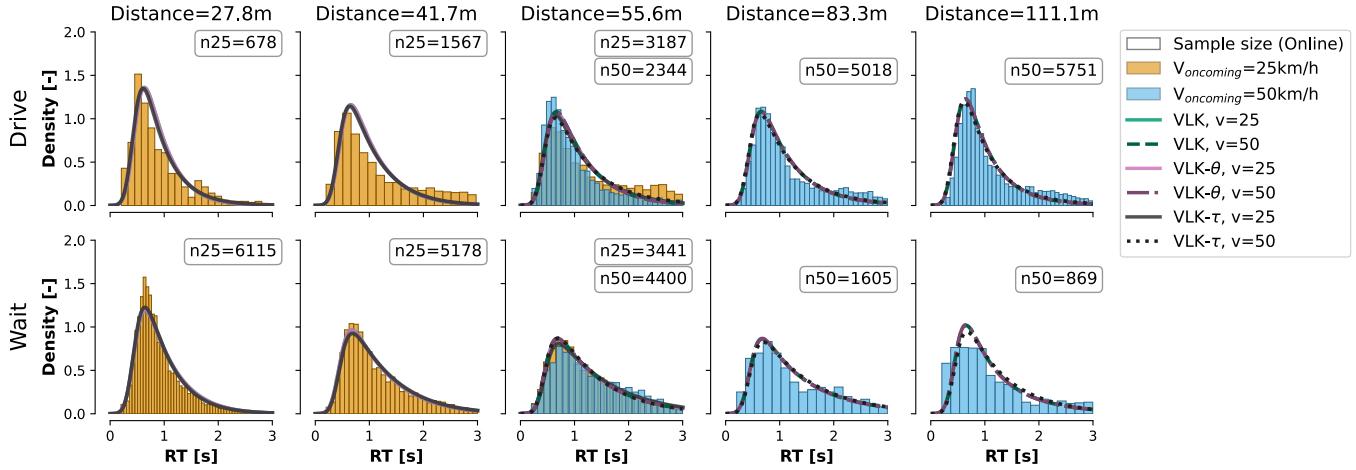
D.4 Mean behavioural measures for lab-trained LK models

Figure 15 and Figure 16 summarise aggregate behavioural measures for the two LK-based model families trained on the full laboratory dataset. Results are shown for the distance-based LK models (DLK; Models 4, 10, and 9) and the velocity-augmented LK models (VLK; Models 0, 18, and 7), respectively.

Each figure reports predicted and empirical values for the probability of choosing to Drive, mean Drive reaction time, and mean Wait reaction time across experimental conditions. Together, these plots provide a compact comparison of model behaviour under controlled laboratory conditions and complement the more detailed online diagnostics presented in Section D.3.



(a) Population-averaged *Drive* probability and mean *Drive* and *Wait* RTs from measured online data and predictions of online-trained velocity-based models, ordered by initial distance between the oncoming vehicle and the narrow passage



(b) Reaction time probability density functions of the online-trained velocity-based DDMs overlaid on online data, ordered by initial distance of the oncoming vehicle to the narrow passage on the horizontal axis and by the decision outcome on the vertical axis. Colours indicate the oncoming vehicle velocity v_{on}. n denotes the number of trials contributing to each distribution and v_{on}.

Figure 14: Extended diagnostic visualisations for online-trained velocity-based DDMs. Condition-level decision and reaction-time summaries (top) are shown alongside distance-ordered reaction-time distributions (bottom), providing a joint view of mean-level and distributional model behaviour under full-data training.

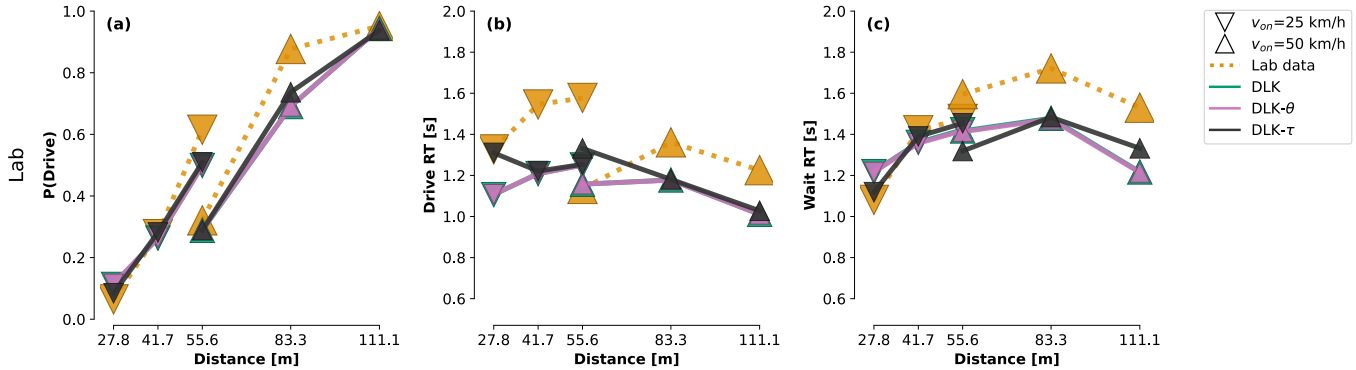


Figure 15: Population-averaged *Drive* probability and mean *Drive* and *Wait* RTs from measured lab data and predictions of lab-trained distance-based models, ordered by initial distance between the oncoming vehicle and the narrow passage

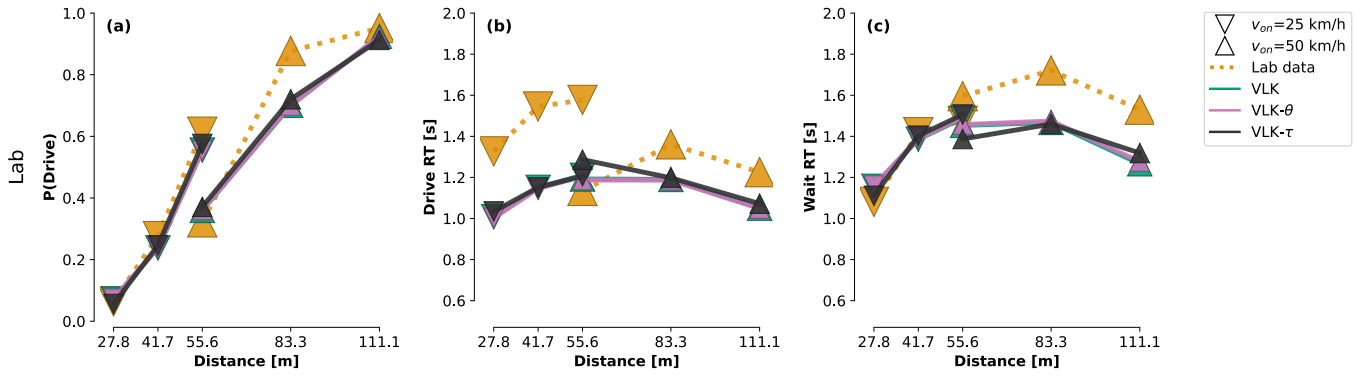


Figure 16: Population-averaged *Drive* probability and mean *Drive* and *Wait* RTs from measured lab data and predictions of lab-trained velocity-based models, ordered by initial distance between the oncoming vehicle and the narrow passage

E Demographic influences and individual differences

This appendix examines the relationship between individual demographic characteristics and behavioural measures in the narrow-passage task. The analyses presented here serve two purposes: first, to assess whether age, driving experience, or habitual driving exposure systematically explain within-dataset variability in decisions or reaction times; and second, to characterise how individual-level patterns differ between the laboratory and online samples. The results are intended to contextualise population-level differences rather than to motivate additional modelling assumptions.

E.1 Within-dataset demographic effects

Figures 17 and 18 show behavioural measures binned by participant age, driving experience, and habitual annual mileage for the laboratory and online datasets, respectively. For each dataset, stacked bars indicate the proportion of Drive and Wait decisions, while paired bars report mean reaction times for Drive and Wait responses.

Across both datasets, no consistent monotonic relationship was observed between any demographic variable and either decision outcome or reaction time. In particular, neither age nor driving experience explained systematic shifts in choice probability or response speed within a given dataset. Similarly, habitual annual mileage showed no reliable association with decisiveness or reaction time once data were aggregated within demographic bins.

These results indicate that demographic variables do not account for the dominant sources of within-dataset behavioural variability observed in the task.

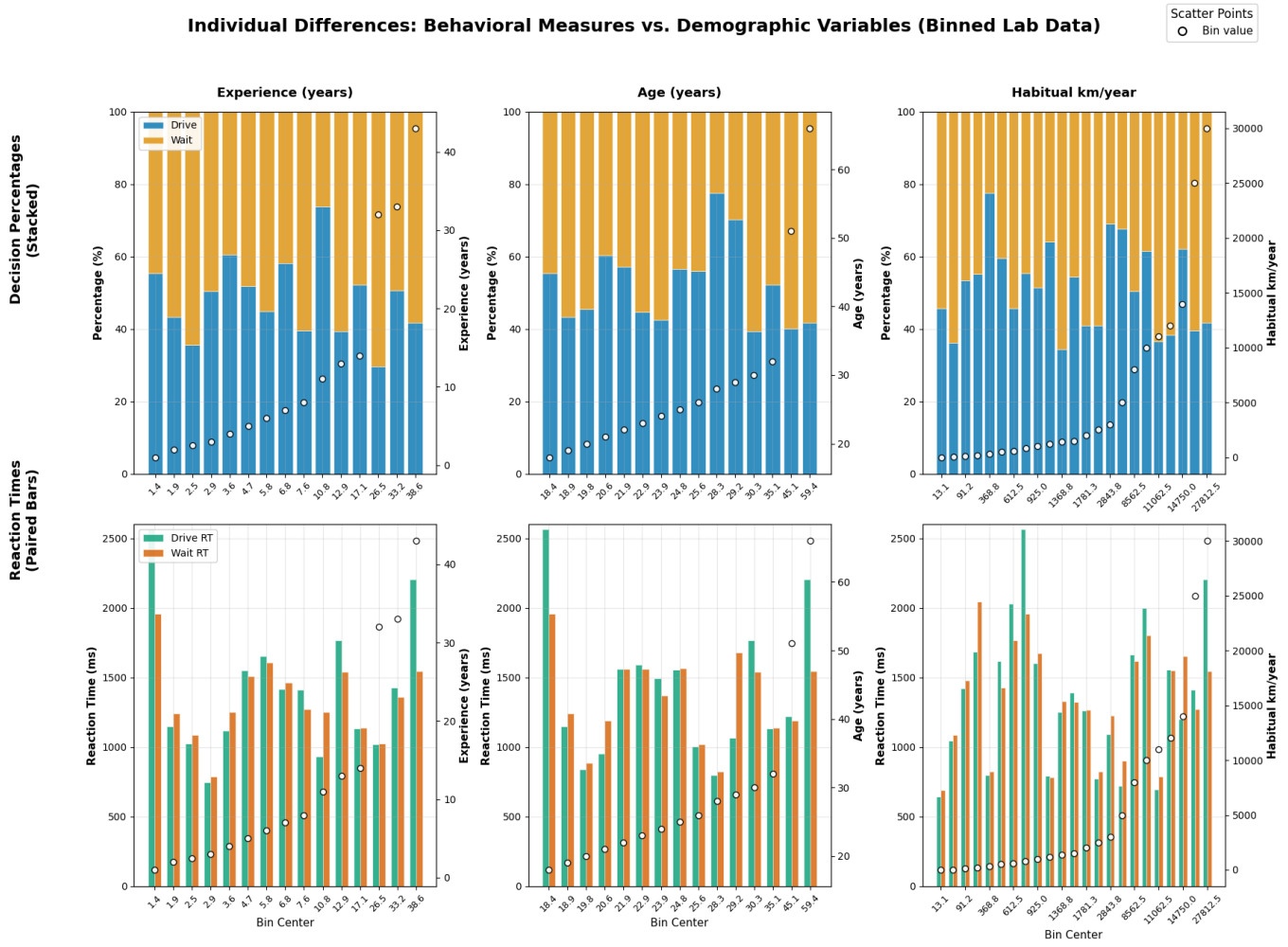


Figure 17: Behavioural measures as a function of demographic variables for the laboratory dataset. Stacked bars show Drive and Wait proportions; paired bars show mean reaction times. No systematic within-dataset associations with age, experience, or habitual mileage are evident.

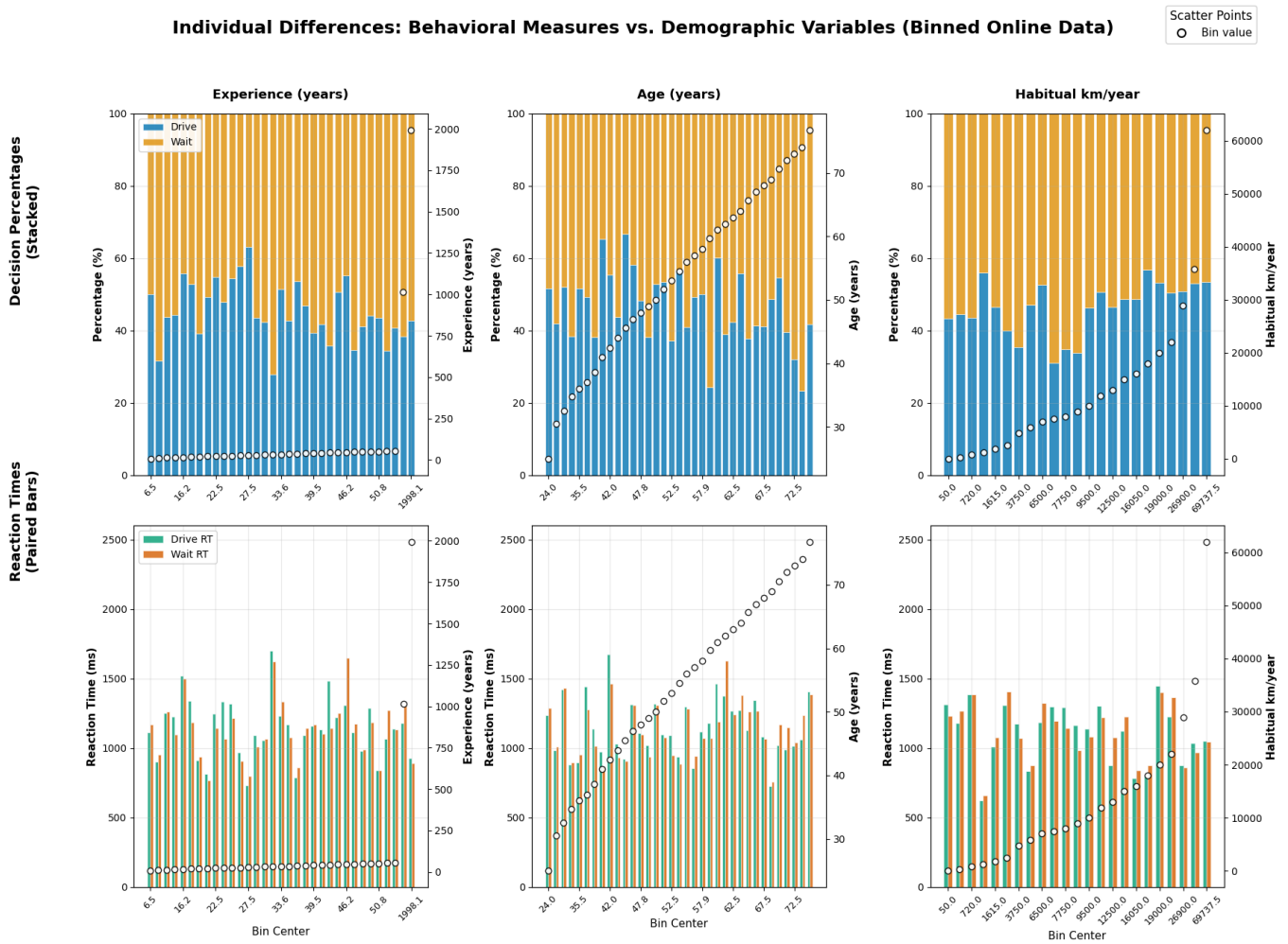


Figure 18: Behavioural measures as a function of demographic variables for the online dataset. Despite a broader age range and increased variability, no consistent within-dataset demographic effects are observed.

E.2 Individual-level patterns across datasets

Figure 19 compares individual participants' mean reaction times and decision proportions across experimental conditions, shown separately for laboratory and online participants. Each panel corresponds to a specific combination of time-to-arrival (TTA_0) and oncoming vehicle velocity v_{on}.

Across most conditions, mean reaction time and decision outcome were only weakly correlated. However, the intermediate-distance conditions–(TTA_0=8 s, v_{on}=25 km/h) and (TTA_0=4 s, v_{on}=50 km/h)—exhibited significant correlations between reaction time and the proportion of Wait responses. In the high-velocity condition, longer reaction times were associated with an increased tendency to Wait, whereas the opposite relationship emerged in the low-velocity condition.

These intermediate conditions are precisely those characterised by increased perceptual ambiguity, as reflected in broadened empirical reaction-time distributions. The observed correlations therefore suggest that, for a subset of participants, additional decision time is used to resolve uncertainty rather than reflecting generic response slowing.

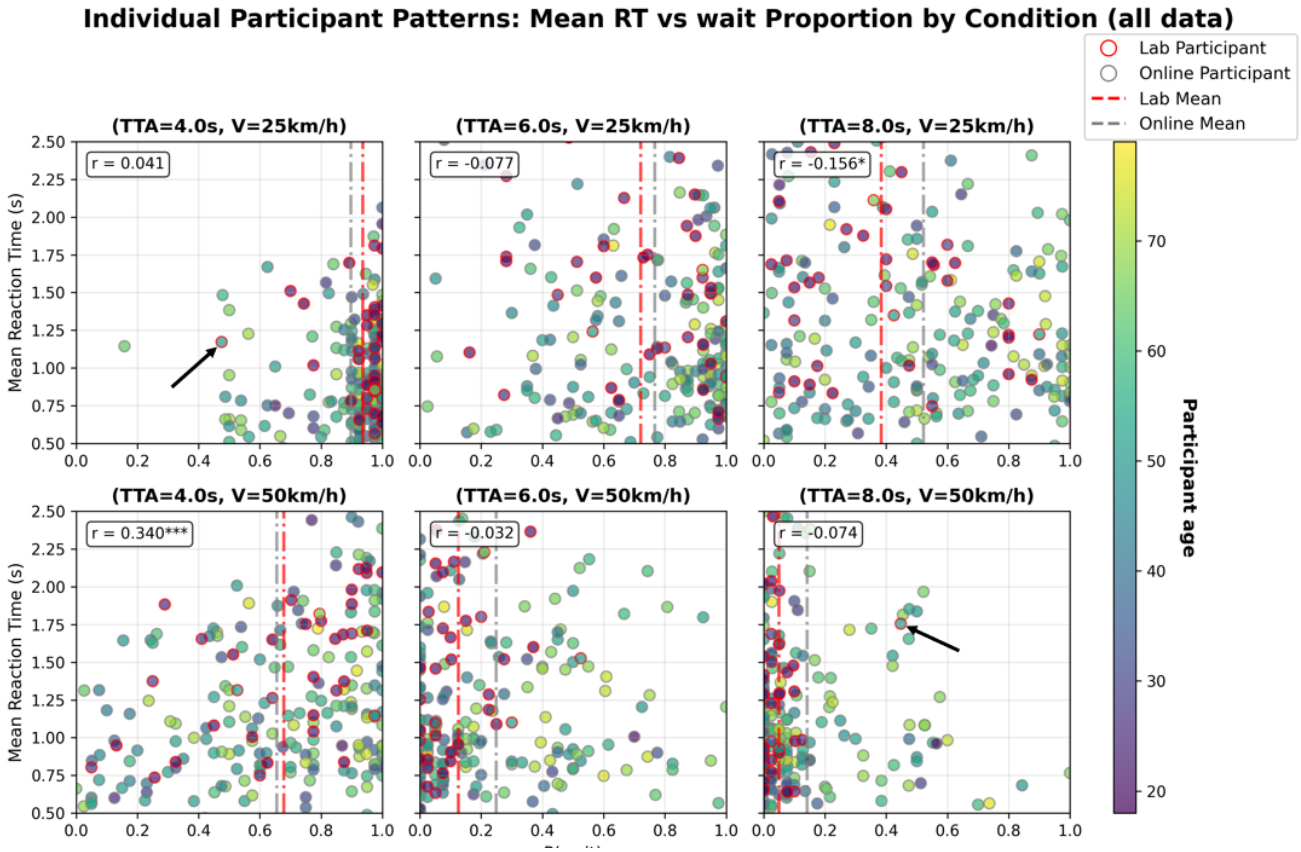


Figure 19: Individual participant patterns relating mean reaction time and decision outcome across conditions. Laboratory and online participants are shown separately. Significant correlations appear only in intermediate-distance conditions, where perceptual ambiguity is highest. The black arrow indicates an exceptionally old lab participant, exhibiting behaviour more in accordance with the online population.

F Cross-dataset and LOCO generalisation

This appendix presents validation analyses assessing the generalisability and robustness of the evidence-accumulation models beyond the conditions under which they were trained. Generalisation is evaluated both across datasets and to unseen experimental conditions using overlay-based comparisons of empirical data and model predictions. All results reported in this appendix are obtained without refitting models to the evaluated data.

F.1 Cross-dataset generalisation

Figure 20 shows cross-dataset generalisation at the level of average behavioural measures, where models trained on the online dataset are evaluated on the laboratory data without refitting. Although absolute choice probabilities and RT means differ between datasets—reflecting differences in sample composition and experimental context—the models reproduce the same qualitative dependence of decision outcome and timing on TTA_0 and oncoming-vehicle velocity v_{on}.

To complement this analysis and to support the corresponding cross-dataset result presented in the main text, Figure 21 reports RT distributions for the inverse transfer direction, where models trained on the laboratory dataset are evaluated on the online data. Presenting the RT-PDFs in this direction allows a direct distributional comparison with the lab-trained overlay shown in the main results. Despite shifts in absolute response times across datasets, the RT-PDFs exhibit closely matched distributional structure across decision outcomes, indicating that the learned timing dynamics generalise across populations.

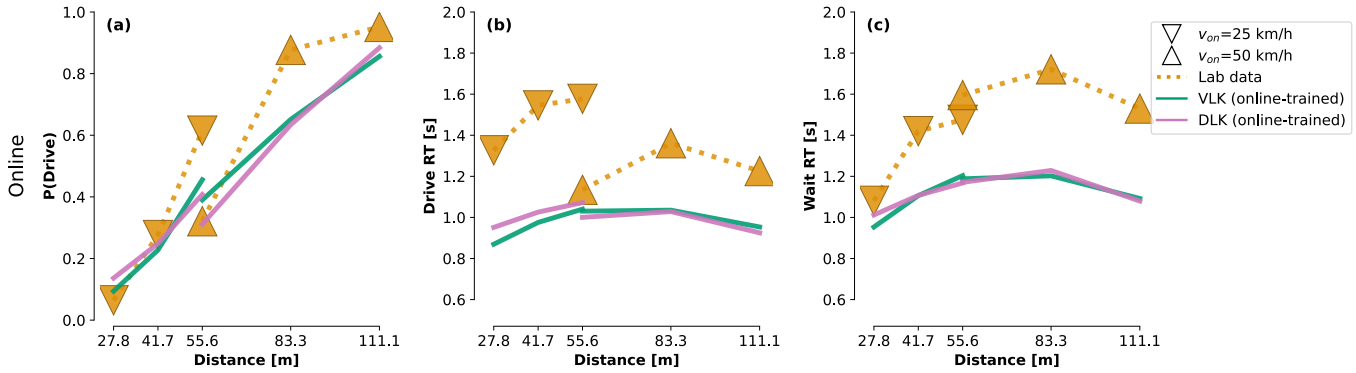


Figure 20: Cross-dataset generalisation from online to laboratory data. Models trained on the online dataset are evaluated on laboratory data without refitting. Panels show P(*Drive*), mean Drive RT, and mean Wait RT. While absolute levels differ across datasets, the qualitative dependence on experimental conditions is preserved.

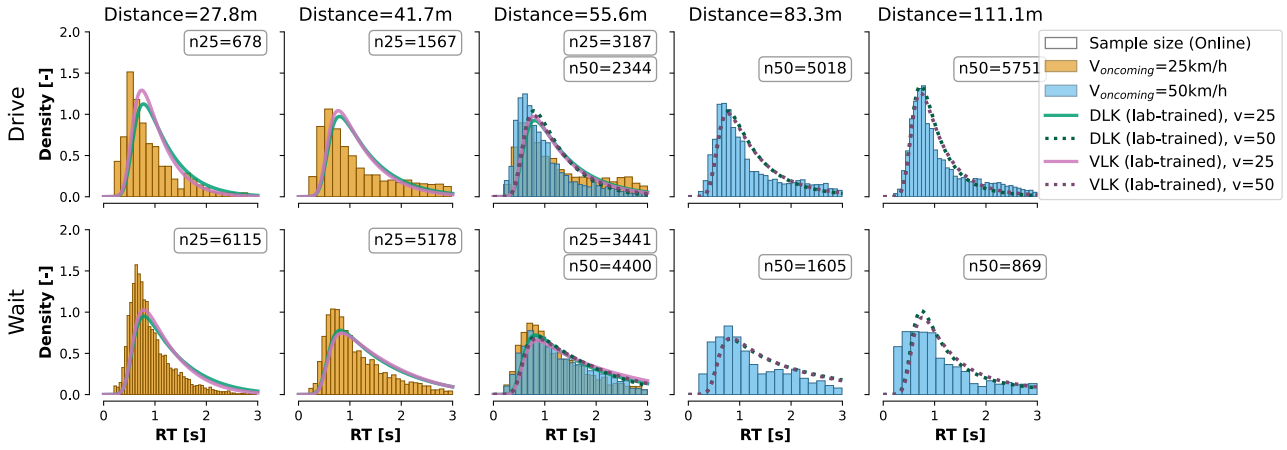
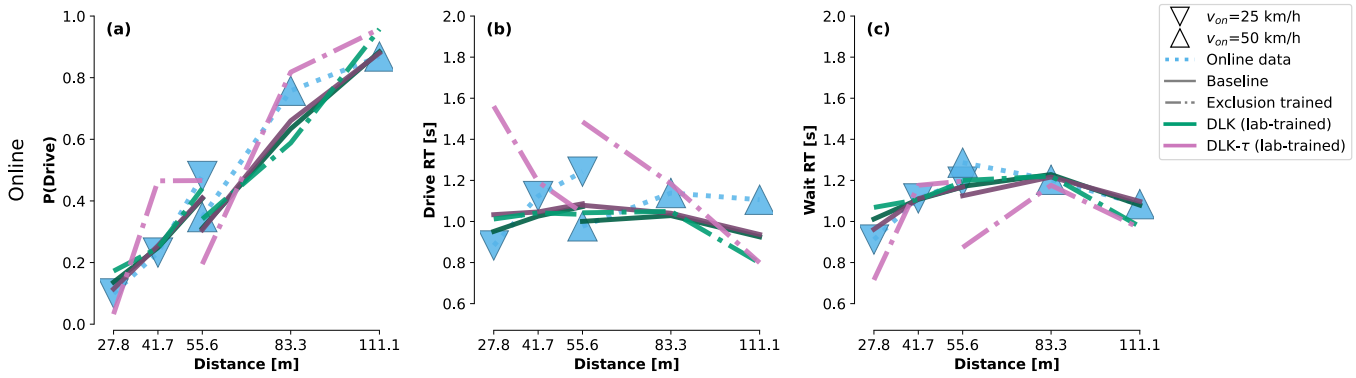


Figure 21: Reaction time probability density functions (RT-PDFs) for cross-dataset generalisation from laboratory to online data. Empirical distributions from the online dataset are overlaid with predictions from models trained on the lab dataset without refitting. Despite dataset-dependent shifts in mean response times, the distributional timing structure remains closely matched across decision outcomes.

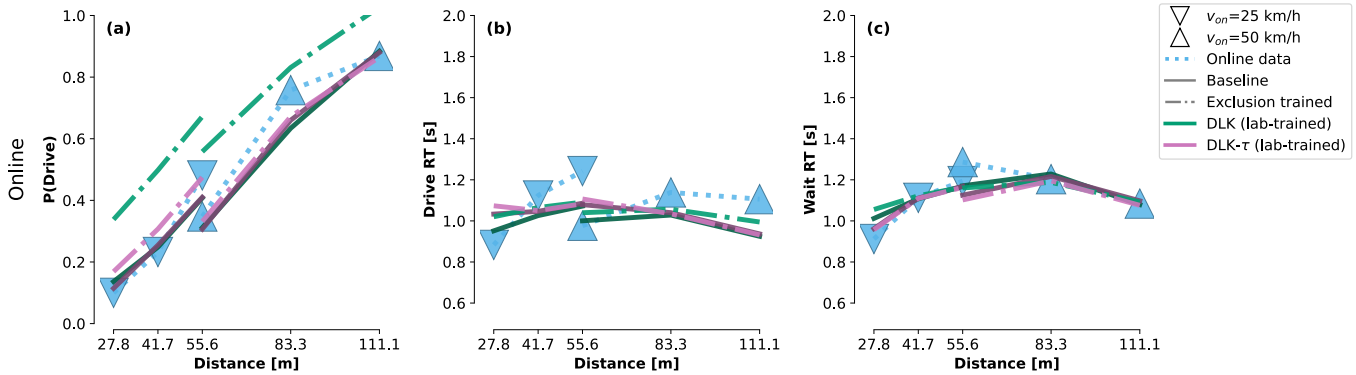
F.2 Leave-one-condition-out training robustness

Worst-case generalisation was assessed using a leave-one-condition-out (LOCO) procedure. For each condition, model predictions were generated using parameters fitted to a training set excluding the condition being plotted.

Figure 22 shows LOCO generalisation for the DLK and DLK-τ models, which represent the most vulnerable model formulation. Results are shown separately for initial-time-to-arrival (TTA_0) exclusions and for single TTA_0×velocity condition exclusions.



(a) LOCO generalisation with TTA-level exclusions.



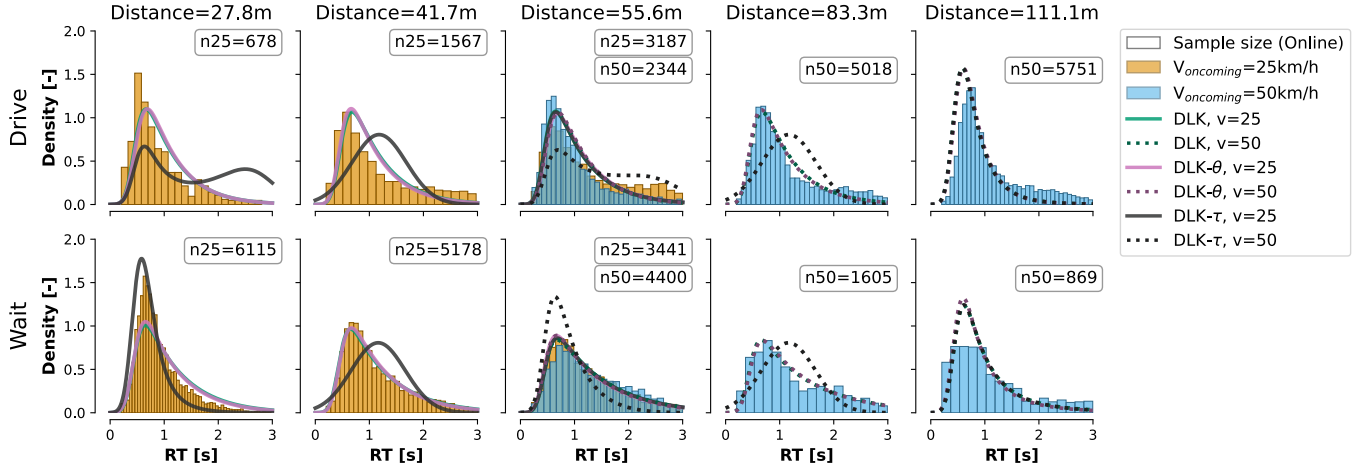
(b) LOCO generalisation with TTA×velocity exclusions.

Figure 22: Aggregate LOCO generalisation for distance-based (DLK) and distance-looming-based (DLK-τ) models. Each point in the plots is predicted by a model trained only on other conditions, so all points are unseen conditions to the models. Panels show P(Drive), mean Drive RT, and mean Wait RT.

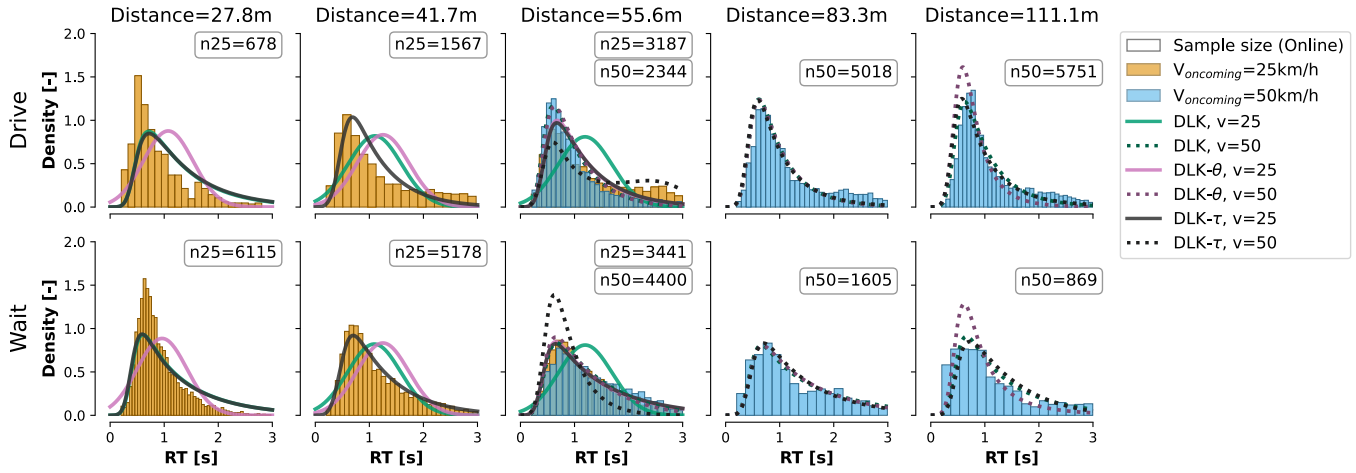
F.3 Stress testing condition exclusion robustness via reaction time probability density functions

To assess whether LOCO generalisation holds at the temporal structure level, RT-PDFs were examined for representative held-out conditions. Figure 23 shows RT-PDFs obtained under TTA and TTA×velocity exclusions, overlaid with predictions from DLK, DLK-θ, and DLK-τ models.

These distributions confirm that the generalisation observed at the level of average measures is not a mean-only artefact, and that timing structure remains qualitatively intact under worst-case exclusion.



(a) RT-PDFs under TTA-level LOCO exclusion.



(b) RT-PDFs under TTA×velocity LOCO exclusion.

Figure 23: Reaction time probability density functions of the online-trained distance-based DDMs with traing data exclusions overlaid on online data, ordered by initial distance of the oncoming vehicle to the narrow passage on the horizontal axis and by the decision outcome on the vertical axis. Colours indicate the oncoming vehicle velocity v_{on}. n denotes the number of trials contributing to each distribution and v_{on}. Each line in the plots is predicted by a model trained only on other conditions, so all points are unseen conditions to the models.

G Model framework generalisation

This appendix evaluates the generality of the longitudinal kinematics drift diffusion modelling (LK-DDM) framework by applying it to a related but distinct interaction scenario, namely the left-turn scenario as described in Appendix B. In contrast to Appendix F, models are refitted here on the complete left-turn dataset, as the objective is to assess framework reusability rather than robustness under fixed parameters.

G.1 Refitted models

Table 11 summarises the reported model-comparison results. Models were compared using the Bayesian Information Criterion (BIC), accounting for both goodness of fit and model complexity. A constant-boundary model with an initial bias and without an explicit acceleration term achieved the lowest BIC and was selected as the best-performing configuration. This model structure closely aligns with the base LK drift formulations considered in the present thesis.

Table 19: Model parameters for DLK and VLK base models refitted to the Left-turn Nudging experiment

model	BIC	α	β_v	β_d	θ	B	x_0	μ_{NDT}	σ_{NDT}
DLK	4611	0.557	—	0.0019	6.420	1.385	0.639	0.149	0.154
VLK	4601	0.704	0.0739	—	7.179	1.414	0.666	0.155	0.160

G.2 Distributional validation on nudging data

To assess whether the refitted LK-based models capture not only aggregate behaviour but also the temporal structure of decisions in the left-turn task, RT distributions were examined. Figure 24 shows empirical RT PDFs overlaid with predictions from the refitted models.

The distributions indicate that the modelling framework reproduces the characteristic shape and spread of reaction times observed in the left-turn experiment. This agreement at the distributional level complements the aggregate behavioural results reported in the main text and supports the conclusion that the LK-DDM framework generalises to a distinct interaction paradigm when refitted.

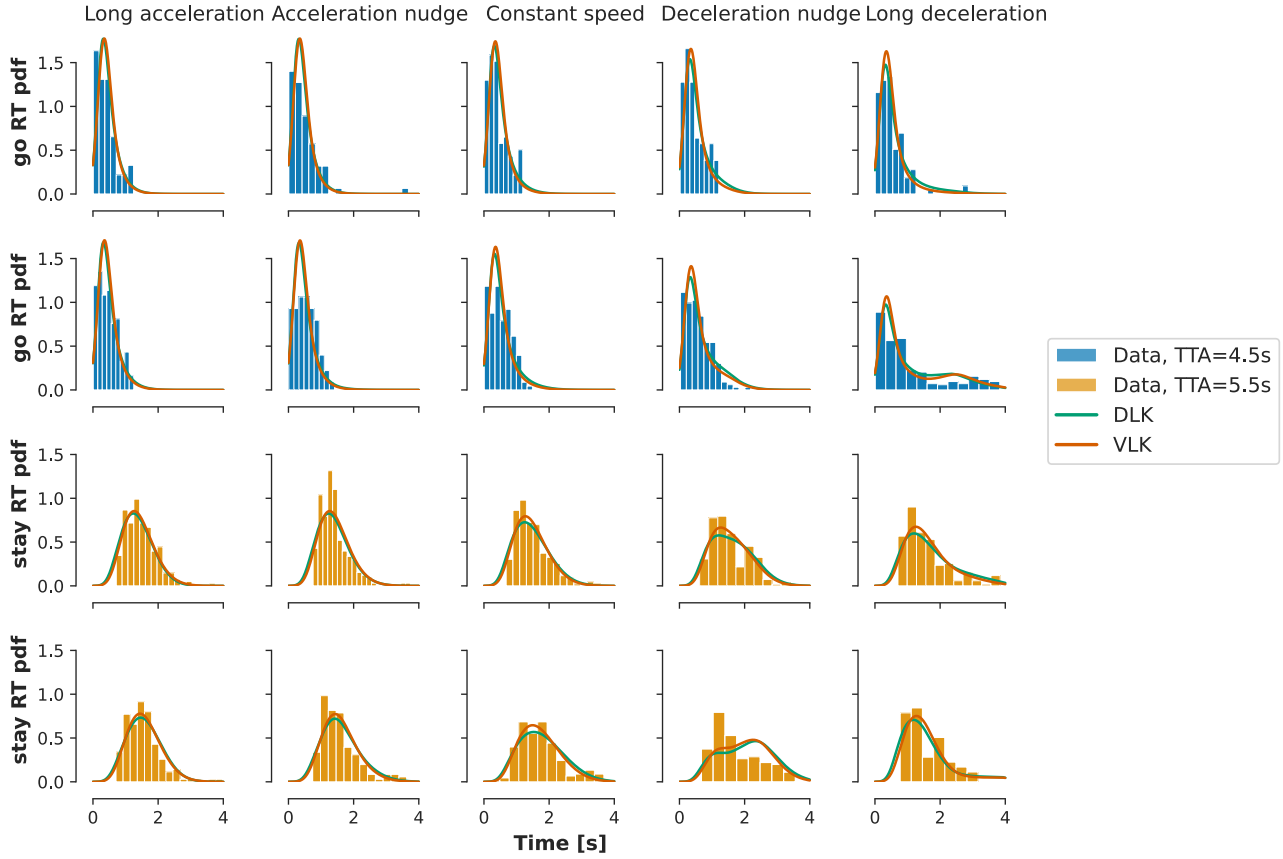


Figure 24: Reaction-time probability density functions (RT-PDFs) for the left-turn nudging experiment. Empirical RT-PDFs are overlaid with predictions from LK-based models refitted to the left-turn dataset. The distributions demonstrate that the modelling framework captures the temporal structure of decisions in a distinct interaction context.