Transportation
https://doi.org/10.1007/s11116-026-10793-z

Perceived traffic safety's impact on cycling route choice with street-level images

Koichi Ito[1,2] · Roos Terra[2] · Miguel Costa[3] · Felix Wilhelm Siebert[3] · Francisco Garrido-Valenzuela[2] · Carlos Lima Azevedo[3] · Filip Biljecki[1,4] · Sander van Cranenburgh[2]

Received: 15 November 2025 / Accepted: 21 June 2026

Abstract

Perceived traffic safety influences cycling behavior, yet its value relative to travel time and other route attributes remains unquantified. Building on the stated choice experiment and street-level image dataset of Terra et al. (2025), we extract safety perception scores from cycling-perspective images using a computer vision model and estimate mixed logit models to test whether perceived safety affects route choice after controlling for visual street-level features. Cyclists are willing to accept 64 additional seconds of travel time for a one-unit increase in perceived safety (scale: -2 very unsafe to $+2$ very safe). Safety preferences vary across demographic groups: older cyclists, recreational cyclists, and those with positive cycling attitudes place more weight on safety, while commuters prioritize speed. These willingness-to-pay estimates enable planners to quantify perceived safety improvements for cost-benefit analyses and to score existing cycling networks for targeted infrastructure upgrades. (Replication code: https://github.com/koito19960406/cycling_safety_perception).

Keywords Cycling route choice · Traffic safety perception · Street-level imagery · Computer vision · Discrete choice modeling

Koichi Ito
koichi.ito@u.nus.edu

1 Department of Architecture, National University of Singapore, 4 Architecture Drive, Singapore 117566, Singapore

2 CityAI Lab, Transport and Logistics Group, Faculty of Technology Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands

3 Department of Technology, Management and Economics, Technical University of Denmark, Akademivej 2800, Kongens Lyngby, Denmark

4 Department of Real Estate, National University of Singapore, 15 Kent Ridge Drive, Singapore 119245, Singapore

Published online: 08 July 2026

Introduction

Cycling is an important component of sustainable urban transport systems, offering environmental, health, and economic benefits across cities worldwide (Brand et al. 2021; Oja et al. 2011; Blondiau et al. 2016). Despite these advantages, cycling adoption remains limited by safety concerns, with perceived risk being the primary barrier to cycling uptake (Useche et al. 2019). Designing infrastructure and policies that encourage cycling therefore depends on understanding cyclists' route choices: how much they value perceived safety relative to the travel time they trade for it, and how that value differs across riders.

The built environment, urban design, and street-level conditions play a central role in shaping how cyclists navigate urban spaces and choose their routes (Lukawska 2024). Unlike motorised transport, cyclists are more exposed to environmental conditions and traffic interactions when making route choices (van Wijnen et al. 1995; Reynolds et al. 2009). Research on cyclists' perceptions has identified that infrastructure characteristics, particularly separation from motor traffic, road surface quality, and interaction with other road users, fundamentally shape how cyclists evaluate routes (Berghöfer and Vollrath 2022, 2023). The deterrent effect of negative route attributes (e.g., heavy traffic, poor surfaces) is generally stronger than the motivational effect of positive attributes, consistent with loss aversion (Berghöfer and Vollrath 2023). Thus, the quality of infrastructure and perceived safety could affect cyclists' route choices. This influence creates both challenges and opportunities for urban planners seeking to promote cycling through targeted infrastructure investments (Meng 2022).

A substantial body of research has examined cyclists' route preferences using stated-preference (SP) methods. Early studies relied primarily on textual descriptions or abstract sketches to represent route attributes, such as the presence of bike lanes, road surface quality, and greenery (Bovy and Bradley 1985; Sener et al. 2009; Vedel et al. 2017). While these approaches provided valuable insights into cyclist preferences, they cannot capture the full complexity and nuance of real-world cycling environments. Text descriptions require respondents to mentally reconstruct scenarios, which can lead to misinterpretation, whereas simplified sketches omit important contextual details, such as surface texture, vegetation quality, and the overall sense of safety that emerges from the complete streetscape. Another stream of research uses revealed preference (RP) data from GPS trackers to model actual route choices (Broach et al. 2012; Ton et al. 2017; Skov-Petersen et al. 2018).

More recently, computer vision and street-level imagery have revolutionised how researchers study urban environments. This research typically focuses on quantifying perceptions of safety, beauty, or "bikeability" (Ito and Biljecki 2021; Costa et al. 2025a; Ito et al. 2024a). While powerful for mapping perceptions across cities, these perception studies do not model choice behaviour itself. Importantly, perceptions (e.g., perceived safety and beauty) might influence preferences only to some extent, given other factors such as travel time. Along with these image-based perception studies, a growing body of research has also emerged building RP route choice models with street-level imagery (Hankey et al. 2021; Ito et al. 2024). Despite the proliferation of RP methods, these methods are limited by existing infrastructure and cannot evaluate preferences for non-existent cycling facilities or capture latent demand from potential cyclists who do not currently cycle. Thus far, only a few existing studies have combined discrete choice SP models with (a large number of) explicitly processed street-level imagery as model inputs (e.g., van Cranenburgh and Garrido-Valenzuela 2025, 2026).

Despite computer vision models' ability to map perceived street safety from street-level images, no study has yet integrated these scores into a stated-choice experiment to quantify how much cyclists value perceived safety improvements relative to other route attributes. While prior work recognises perceived safety as a factor in cycling behaviour (Uijtdewilligen et al. 2024; Kuiper 2021), its magnitude in terms of willingness-to-pay and its variation across demographic groups remain unquantified. Filling this gap would enable urban planners to put a concrete value on improvements in safety perception and inform infrastructure investments accordingly.

This study bridges this gap by combining large-scale street-level imagery with discrete choice modeling. Cycling route choice typically involves trade-offs among multiple route attributes (e.g., travel time, the number of traffic lights at intersections, comfort, perceived safety, and others). Our analysis quantifies how cyclists trade perceived safety against the two attributes that vary in our stated-preference experiment: travel time and the number of traffic lights. We use a dataset from the stated choice experiment by Terra et al. (2025), which uses thousands of real street-level images of Rotterdam captured from the perspectives of cyclists, providing a more accurate representation than the vehicle viewpoint imagery used in mainstream platforms like Google Street View (Ito et al. 2024b). We then derive perceived safety scores from these images with a recently developed computer vision model by Costa et al. (2025a) and estimate mixed logit discrete choice models to examine whether perceived safety has a significant effect on cyclists' route choices after controlling for other visual features, such as greenery and cars (which we extract using semantic segmentation). Thereby, our work addresses three research questions:

RQ1: Does perceived traffic safety affect cyclists' route choices after controlling for visual street-level features (e.g., vegetation, bike lanes, and vehicles) extracted through image segmentation?

RQ2: What is cyclists' willingness-to-pay for perceived safety improvements, expressed in additional travel time and traffic lights?

RQ3: To what extent do preferences for perceived traffic safety vary across demographic groups, and which groups attach relatively greater or lower importance to safer cycling routes?

The remaining part of this paper is organized as follows. Literature review section reviews the literature on cycling route choice modeling and computer vision applications in transport research. Data and methods section describes our data collection and methodological approach, including the stated choice experiment design, safety perception model, and discrete choice modeling framework. Results section presents the results from our main models and interaction analyses. Discussion section discusses the key findings and their implications for urban planning. Conclusion section concludes with limitations and directions for future research.

Literature review

This review traces what is known about cyclists' route preferences, focusing on how far prior work has gone in quantifying the value cyclists place on perceived safety.

Evolution of cycling route choice modeling

Cycling route choice modeling has evolved considerably over the past three decades, driven by advances in survey methods, modeling techniques, and data availability. The field broadly divides into two methodological approaches: revealed preference (RP) and stated preference (SP) studies, each offering distinct advantages and limitations for understanding cyclist behavior.

RP studies use observational data, such as GPS tracks or smartphone data, to model the routes cyclists actually take (Broach et al. 2012; Skov-Petersen et al. 2018; Ton et al. 2017; Zhang et al. 2021; Xiao et al. 2025). These studies consistently identify travel distance as the dominant factor in route choice, though the influence of infrastructure varies significantly by urban context. While RP studies capture actual behavior, they face inherent limitations in disentangling confounding factors and quantifying trade-offs. Chosen routes may not represent clear alternatives, and there is a potential bias introduced when extracting street attributes by matching GPS trajectories to street segments. RP data also cannot easily quantify how much additional travel time cyclists would accept for infrastructure improvements, a metric needed for cost-benefit analysis of cycling investments. The concept of worthwhile travel time, where travel conditions make additional journey time acceptable or even desirable, has been reviewed and synthesized in transport research (Cornet et al. 2022). However, while a few cycling studies have estimated travel time trade-offs for infrastructure features such as bike paths and surface quality (Broach et al. 2012; Berghöfer and Vollrath 2023), the accepted travel time cost of perceived safety improvements remains poorly quantified, particularly using image-based safety measures.

SP studies can overcome these limitations by presenting respondents with carefully controlled hypothetical choice scenarios. This systematic variation of attributes enables direct estimation of trade-offs and willingness-to-pay values. Early SP studies relied heavily on textual descriptions to convey route attributes such as “road width,” “amount of greenery,” or safety levels (Sener et al. 2009; Stinson and Bhat 2003; Rossetti and Daziano 2023). While these studies established fundamental patterns in cyclist preferences, they required respondents to mentally reconstruct complex scenarios, potentially introducing bias through misinterpretation or incomplete visualization (Hevia-Koch and Ladenburg 2019).

Recognizing these limitations, later studies incorporated visual elements to improve realism. Some studies have used short video clips Tilahun et al. (2007); Griswold et al. (2018), while others employed sketches (Vedel et al. 2017) or abstract edited images (Hardinghaus and Papantoniou 2020; Rossetti et al. 2018). However, even these enhanced approaches face limitations. Simplified visuals and edited images still lack the richness of real-world environments and may not elicit the same responses as actual streetscapes. Abstract representations often omit crucial contextual details such as surface texture, lighting conditions, and the overall environmental quality that influences cycling comfort and safety perceptions. Moreover, these studies did not quantify the perceptual qualities (e.g., perceived cycling safety) directly and explicitly from presented images to measure their effects.

Cycling safety research distinguishes between *objective safety* (measured through crash statistics, infrastructure audits, and traffic conflict analysis) and *perceived safety* (the subjective assessment of risk based on environmental cues and personal experience). The relationship between the two is complex and context-dependent, with studies reporting positive alignments, negative correlations, and no associations depending on the measures and settings used (Costa et al. 2025b). For instance, Winters et al. (2012) found a moderate correlation ($r = -0.68$) between perceived and observed cycling risk across 14 route types, yet with notable discrepancies for specific infrastructure types. Of the 54 road safety measures reviewed by Sorensen and Mosslemi (2009), 25 had opposite effects on objective and subjective safety. Objective scoring methods and subjective ratings cannot reliably predict actual crash occurrence either (Fuest et al. 2023). Perceived safety is therefore an independent construct that warrants separate investigation, not a simple proxy for objective safety. In a revealed preference study, Huber et al. (2024) reported that perceived safety, measured through survey-reported incidents, did not significantly influence actual route choices, which may reflect either that SP designs surface safety preferences suppressed under the constraints of observed route choice, or that RP safety indicators (self-reported incidents) and CV-derived perception scores measure different constructs. The relative reliability of the two approaches for recovering safety preferences remains open.

Emerging role of computer vision in urban transport research

The proliferation of street-level imagery and computer vision has facilitated novel approaches for studying urban environments and transport behavior. This technological advancement has enabled researchers to analyze visual environments at scale while capturing details that traditional survey methods cannot adequately represent (Biljecki and Ito 2021).

Computer vision techniques have been increasingly applied to quantify subjective perceptions of urban environments using crowd-sourced assessments. Studies typically employ pairwise image comparisons where participants evaluate attributes like safety, beauty, or walkability (Dubey et al. 2016; Salesses et al. 2013). These human perception data are then used to train machine learning models that can predict perceptual scores for new images across entire cities (Wang et al. 2024). The resulting perception maps provide valuable insights into spatial patterns of environmental quality and help identify areas needing improvement (Ogawa et al. 2024).

Research has specifically applied these techniques to cycling environments successfully. Ito and Biljecki (2021) developed a bikeability index using street view images from multiple cities, finding that street-level visual indicators often correlate more strongly with cycling attractiveness than macro-level metrics. Costa et al. (2025a) created perception models for cycling safety using pairwise image comparisons, revealing how visual elements contribute to safety perceptions across different urban contexts. Their model employed a Siamese-style convolutional neural network architecture with dual sub-networks for both classification and ranking tasks, trained on thousands of pairwise image comparisons from Berlin. The model achieved strong performance by combining binary/ternary (i.e., win, lose, or tie) classification loss with a ranking loss that enforces meaningful margins between safer and less safe cycling environments. There is also a growing body of studies correlating visual features with RP data (Ito et al. 2024). The most proliferated method is street- or point-level aggregation, where researchers aggregate the cycling volume at certain points or street segments and join them with visual features from street-level images to run regression analysis (Ito et al. 2024; Hankey et al. 2021; Zhang et al. 2025; Song et al. 2024; Gao and Fang 2025). And other methods include comparisons between actual and shortest paths (Verhoeven et al. 2018) and a reinforcement learning-based method with GPS trajectory data (Ren et al. 2024). Despite such proliferation, these RP-based studies cannot evaluate preferences for route alternatives that do not exist or analyze the preferences among the currently non-cyclists.

Most existing studies rely on imagery from mainstream platforms like Google Street View, which are captured from car perspectives and may introduce systematic biases when analyzing cycling environments (Ito et al. 2024b). Google's terms of service also restrict computational processing of Street View imagery, limiting its use in automated research pipelines. These car-captured images often do not cover dedicated cycling infrastructure that vehicles cannot access and may not accurately represent the visual experience of cyclists due to differences in viewpoint height and perspective (Fan et al. 2025).

Recently, there has been an emergence of stated-choice experiments that utilize street-level images, which demonstrate the possibility of filling the gaps identified so far. For example, van Cranenburgh and Garrido-Valenzuela (2025) is one of the few studies that combined SP discrete choice models with street view image data in the context of residential location choice, and (Terra et al. 2025) also used a similar research design to test the hypothesis of whether visual information plays a significant role in cyclists' route choices. However, no studies have combined computer vision-derived perceived safety scores with stated preference models to estimate willingness-to-pay for safety improvements in cycling route choice.

Integration challenges and methodological gaps

Despite this progress in both behavioral choice modeling and computer vision, a basic behavioral quantity remains unestablished: how much cyclists value perceived safety, and how that value varies across riders. No prior study brings the two literatures together in a way that can recover it. Table 1 summarises recent studies. Yu et al. (2024) and Ye et al. (2024) use street-level imagery to score perceived cycling safety across the network, but stop at producing perception maps and do not estimate behavioral parameters in a route-choice model. Lu et al. (2025) take the opposite step and embed image-derived built-environment features inside a revealed-preference route-choice model, but without an explicit perceived-safety score and without willingness-to-pay estimation. The behavioral modeling tradition has examined safety-related route attributes using revealed preference (Broach et al. 2012; Kuiper 2021; Huber et al. 2024) or stated preference (Uijtdewilligen et al. 2024) data, but with proxies (infrastructure type, crash counts, self-reported incidents) rather than continuous CV-derived perception scores. Our study sits at the intersection: we incorporate a CV-derived perceived-safety score as a continuous explanatory variable inside a discrete choice model estimated on stated-preference data, estimate willingness-to-pay directly in WTP space, and analyze systematic heterogeneity across 17 demographic dimensions. This combination quantifies how much cyclists value perceived-safety improvements and how those valuations vary across the cycling population.

Data and methods

This section describes the data and our methodological approach. Figure 1 provides an overview of the method, which integrates a large-scale image dataset into a stated choice experiment; extracts features from the images with semantic segmentation, safety score

Table 1 Comparison of recent studies examining cycling safety and route choice

Study	Choice model	Perception scoring	Proxy for safety	Trade-off measure	Heterogeneity	Model
Broach et al. (2012)	RP	No	Infrastructure type	Equiv. % distance	Commute/non-commute	PSL
Kuiper (2021)	RP	No	Crash + infrastructure	No	Age, gender, cycling exp	MNL-PS
Huber et al. (2024)	RP	No	Crash + incidents	No	No	MNL
Uijtdewilligen et al. (2024)	SP	No	Route attributes	No	Age, gender, commuter	Mixed logit
Yu et al. (2024)	–	Yes	Perception (ViT) + crash	No	No	ViT + LightGBM
Ye et al. (2024)	–	Yes	Crash-based (CSL-RE)	No	No	XGBoost + SHAP
Lu et al. (2025)	RP	No	Built environment (SLI)	WTC (% distance)	No	PSL
Our study	SP	Yes	CV safety score	WTP (time, TL)	17 demographics	Mixed logit

WTP, willingness-to-pay; WTC, willingness-to-cycle; PSL, path size logit; MNL, multinomial logit

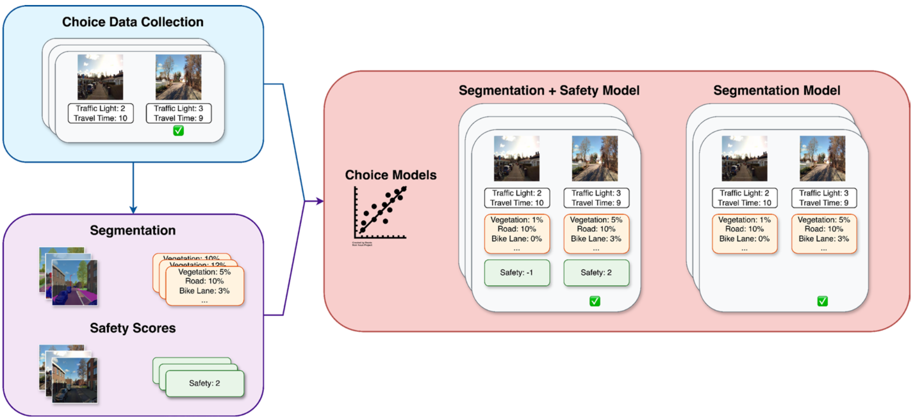

Fig. 1 Overview of the methodological approach combining street-level imagery, computer vision, and discrete choice modeling

prediction, and image embedding; and then compares the model fit among different model specifications.

Datasets

This research builds on the dataset and stated choice experiment by Terra et al. (2025), which combines a collection of geo-referenced street-level images captured from cycling perspectives with stated choice data from a cycling route choice experiment. Focusing on commuting trips, respondents were recruited through a panel in the Netherlands in 2024 and asked to choose between two cycling route alternatives across 15 choice tasks per respondent. In each task, participants selected their preferred route based on the image in combination with other route attributes. The sample was stratified by age and gender to approximate the Dutch cycling population.

The process involved several stages:

Data curation & stated choice experiment: We use the images and choice data collected and curated by Terra et al. (2025). In these data, respondents were asked to choose between two alternative routes. Each alternative was defined by three attributes: travel time (8, 11, or 14 min), number of traffic lights (1, 2, or 3), and the cycling environment, represented by a real street-level image randomly selected from the curated database based on its category. Each task displayed the two alternatives side by side, with a street-level image at the top of each alternative and its travel time and number of traffic lights shown directly beneath, so that respondents weighed the cycling environment together with the numeric attributes rather than in isolation. Figure 2 shows an example choice task.

Beyond the route alternatives, the data set also comprises contextual attributes for each image. These attributes include road type (e.g., separated cycle path, normal road, painted cycle path, access road, main road) and are retrieved from GIS databases. In total, 752 participants completed the survey, resulting in 11,289 valid choice observations. After data cleaning (removing the duplicated tail rows for one respondent and six respondents with fewer than 15 valid choice tasks), the analysis sample comprises 746 participants and 11,190 choice observations. Table 2 provides further summary statistics of the dataset. Readers interested in more details on the data collection are referred to Terra et al. (2025).

Table 3 summarises the sample demographics. The sample is roughly balanced by gender (50.3% female, 49.5% male) and spread across age groups, with slightly fewer respondents aged 71+ (9.0%). Nearly half (48.8%) use the bicycle as their primary transport mode, and most cycle at least twice per week (80.8%). Education levels range from primary/secondary

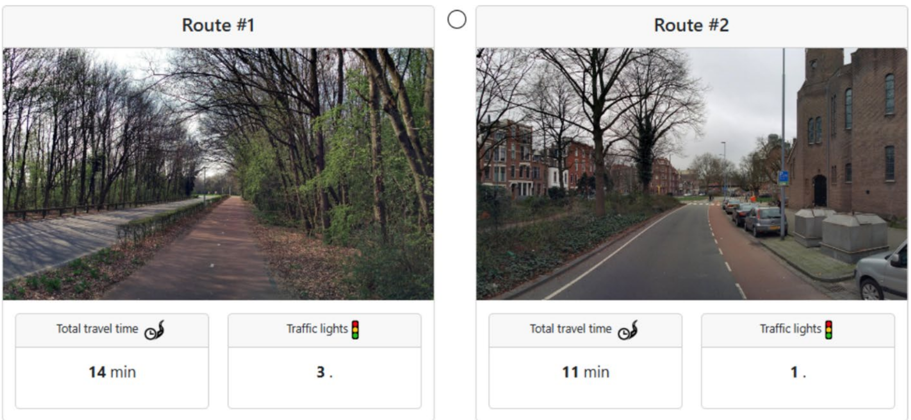

Fig. 2 Example choice task from the stated choice experiment of Terra et al. (2025). Each of the two route alternatives is presented as a street-level image with its travel time and number of traffic lights shown beneath; respondents selected their preferred route

Table 2 Summary statistics of the analysis sample and the choice-task attributes

	Mean	SD	Min.	Max.
Sample				
Number of participants	746			
Number of choice tasks	11,190			
Number of unique images	6,343			
Choice-task attributes (per alternative)				
Travel time, min (8 / 11 / 14)	11.05	2.36	8.00	14.00
Traffic lights, count (1 / 2 / 3)	2.02	0.81	1.00	3.00
Perceived safety score	0.24	0.60	−2.01	2.57

Attribute statistics pool both route alternatives across the 11,190 choice tasks ($N = 22{,}380$ alternatives). Travel time and the number of traffic lights were experimentally varied at fixed levels (shown in parentheses); the perceived safety score is the continuous output of the computer vision model, whose rating anchors range from -2 (very unsafe) to $+2$ (very safe)

Table 3 Descriptive statistics of the survey sample ($n = 746$)

Variable	Category	n	%
Age	18–30	150	20.1
	31–45	167	22.4
	46–60	195	26.1
	61–70	167	22.4
	71+	67	9.0
Gender	Male	369	49.5
	Female	375	50.3
	Other	2	0.3
Household size	1 person	185	24.8
	2 people	276	37.0
	3 people	123	16.5
	4 people	111	14.9
	5 people	36	4.8
	6+ people	15	2.0
Primary transport	Bicycle	364	48.8
	Car	259	34.7
	Walking	75	10.1
	Public transport	41	5.5
	Other	7	0.9
Cycling frequency	5+ days/week	250	33.5
	3 days/week	137	18.4
	2 days/week	113	15.1
	4 days/week	103	13.8
	1 day/week	68	9.1
	<1 day/week	75	10.1
	Never	0	0.0
Education (grouped)	Primary/secondary	98	13.1
	Vocational	327	43.8
	University (B.Sc./B.A.)	186	24.9
	Postgraduate	127	17.0
	Other/prefer not to say	8	1.1

(13.1%) to postgraduate (17.0%), with vocational education as the largest group (43.8%). An example of the raw response data is provided in Table 10 in the Appendix.

Image preprocessing: For model estimation, two additional sets of features are generated. First, semantic segmentation is performed on all images using Mask2Former Cheng et al. (2022a) to quantify the pixel percentage of various urban features. Second, a perceived traffic safety score is computed for each image using a cycling safety perception model developed by Costa et al. (2025a).

Perceived safety model validation: To validate the applicability of the safety perception model trained on Danish and German contexts to the Dutch context, we examine the predicted perceived safety scores across different road types and land use categories. The perceived safety score distributions in Figs. 3 and 4 show clear patterns across different urban contexts. Scores are generally higher for separated cycling paths compared to other road types, while normal road and painted cycle paths did not show much difference. This confirms the importance of physical separation from traffic for perceived safety. Among land use types, recreational areas and main roads show higher safety scores, while industrial areas, residential areas, and access roads show lower scores. These patterns align with expected relationships between urban environment characteristics and cycling safety perceptions, supporting the model's validity in the Rotterdam context.

Figures 5 and 6 show example images scoring minimum, median, and maximum perceived safety scores. For each image, we present semantic segmentation visualizations (showing different urban elements in distinct colors based on the Mapillary Vistas color scheme), and Gradient-weighted Class Activation Mapping (Grad-CAM) overlays (Selvaraju et al. 2016). Grad-CAM is an interpretability technique that produces heatmaps showing which image regions most influenced a neural network's prediction. It backpropagates gradients from the output to the last convolutional layer and weights each activation map by its gradient, producing a spatial relevance map. In our visualizations, higher-intensity colors indicate regions with greater influence on the predicted safety score. The model assigns higher perceived safety scores to environments with better visibility, separation from traf-

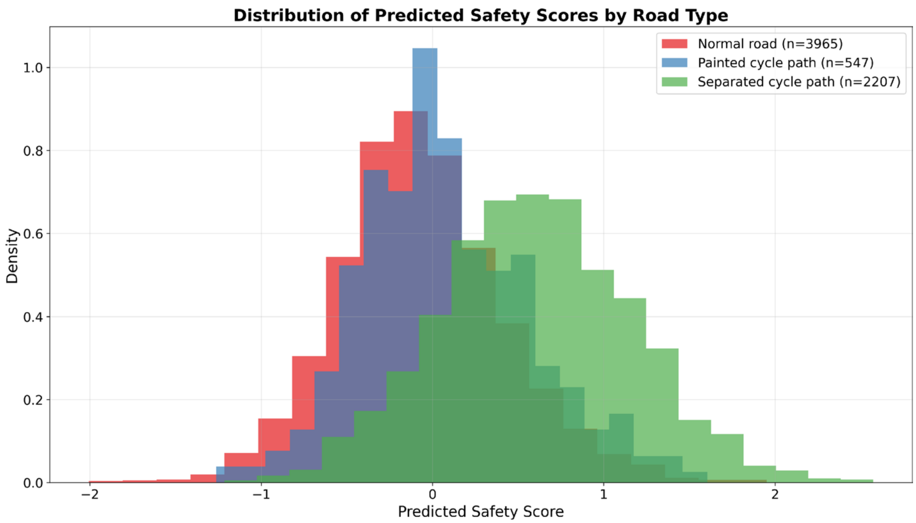

Fig. 3 Distribution of perceived safety scores across different road types

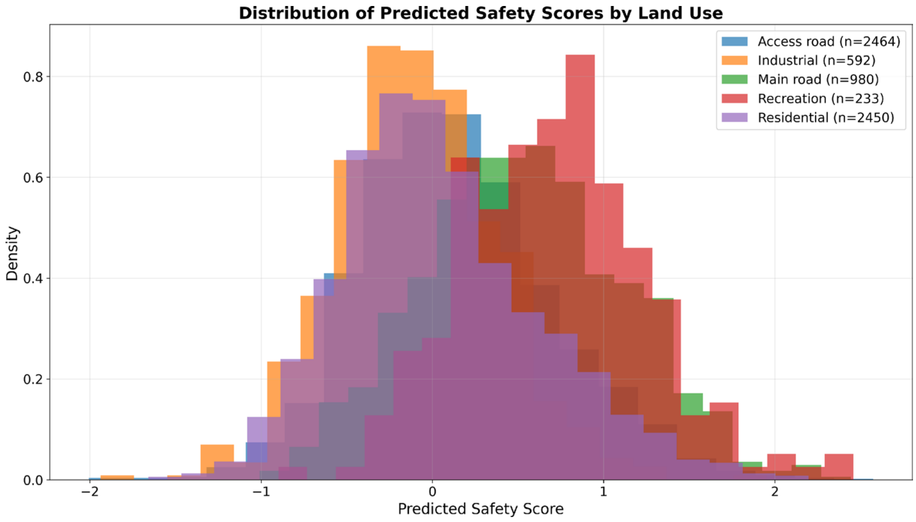

Fig. 4 Distribution of perception of safety scores across different land use categories

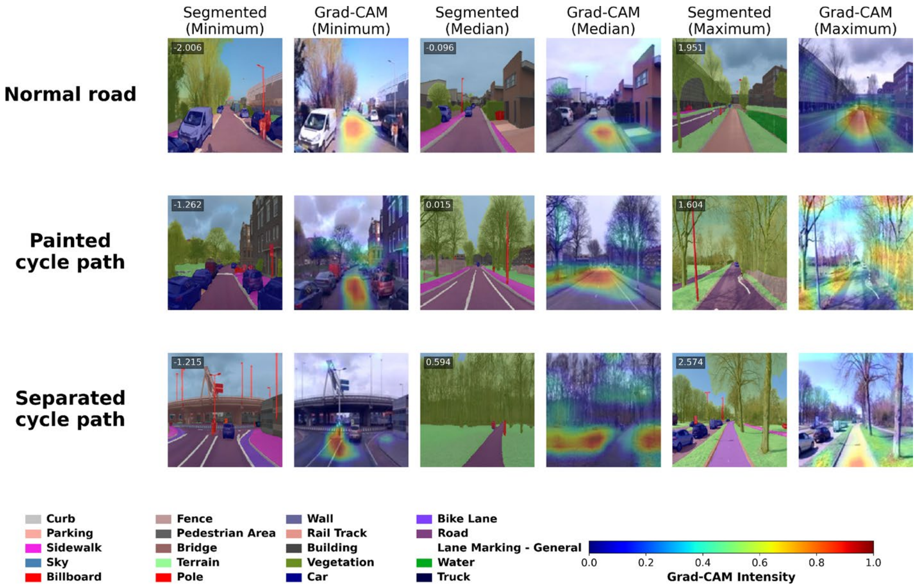

Fig. 5 Example images of predicted safety perception by road type, with attention maps

fic, and well-maintained road conditions. Analysis of attention maps reveals that the model focuses on different environmental features depending on the infrastructure context, paying more attention to road surfaces for normal roads and surrounding elements like vegetation for painted cycle paths.

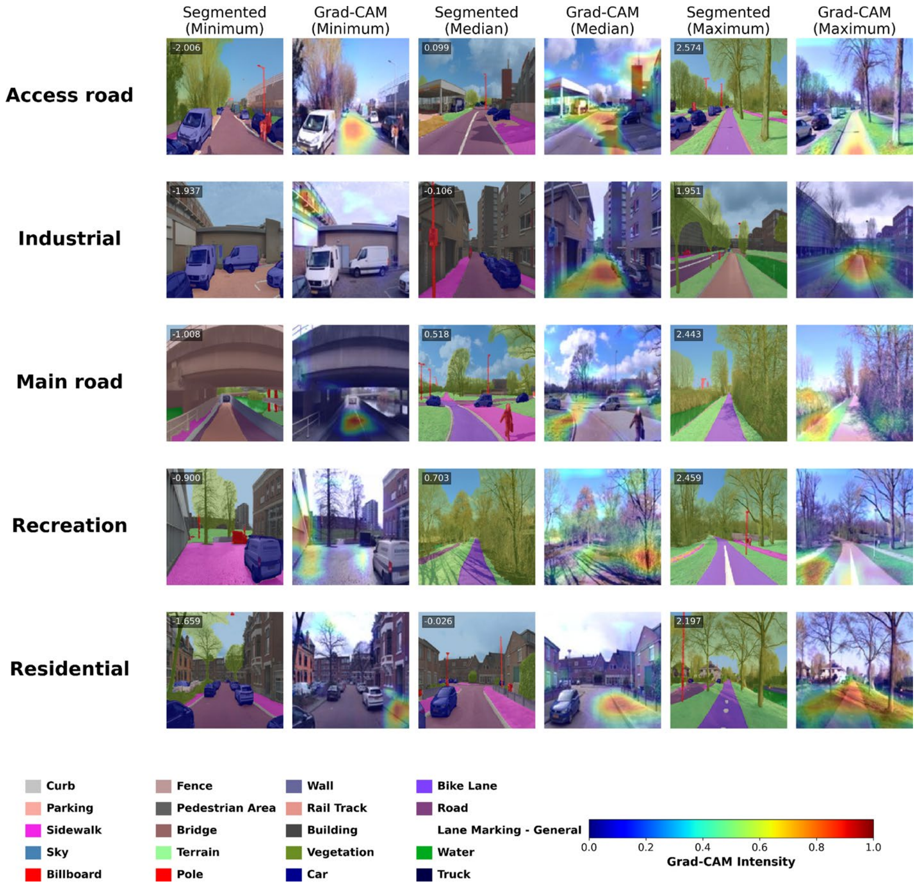

Fig. 6 Example images of predicted perceived safety perception by land use type, with attention maps

Methods

Our methodological approach follows a systematic process where each route alternative in the choice experiment is characterized by three key attributes: travel time (TT), number of traffic lights (TL), and street-level cycling environment represented by real images. To extract comprehensive information from the images, we apply two computer vision models: semantic segmentation to quantify urban environmental features and a safety perception model to derive perceived safety scores. We then estimate mixed logit discrete choice models using different combinations of these features to examine the role of perceived safety in cycling route choice behavior. The replication code can be found here: https://github.com/koito19960406/cycling_safety_perception.

Semantic segmentation

Prior research has established that visual street-level features influence cycling route preferences: greenery is associated with higher route attractiveness (Lu et al. 2025), the presence of motor vehicles deters cyclists (Broach et al. 2012; Uijtdewilligen et al. 2024), and dedicated cycling infrastructure increases route utility (Kuiper 2021). To control for these known effects and isolate the residual contribution of perceived safety, we quantify the pixel-level composition of each image using semantic segmentation.

We apply the Mask2Former model (Cheng et al. 2022b), pre-trained on Mapillary Vistas, a large-scale street-level image dataset with pixel-level annotations across 65 object classes spanning road infrastructure, nature, vehicles, buildings, and construction elements. The model achieves 64.7% mIoU on this dataset. We chose semantic segmentation over alternative approaches (instance segmentation, panoptic segmentation) because it produces pixel proportions for each scene component, which directly serve as continuous explanatory variables in our utility functions. Instance segmentation identifies and counts discrete objects but discards uncountable scene elements (e.g., road surface, sidewalk, vegetation, sky) that are critical for cycling safety perception. Panoptic segmentation combines both instance and semantic approaches but adds complexity without clear benefit for our modeling purpose (Ito et al. 2025).

Traffic safety perception model

We leveraged a model developed by Costa et al. (2025a) to generate a quantitative measure of perceived safety for each image. This model was trained on a large dataset of pairwise image comparisons from Berlin. For our study, we apply this pre-trained model to our Rotterdam image dataset to obtain a continuous perceived safety score for each image in our choice experiment.

The model's architecture, shown in Fig. 7, is a Siamese-style convolutional neural network (CNN) designed to process pairs of images simultaneously. The architecture consists of a shared CNN backbone that extracts feature embeddings from each image, followed by two parallel sub-networks: a classification sub-network and a ranking sub-network. This

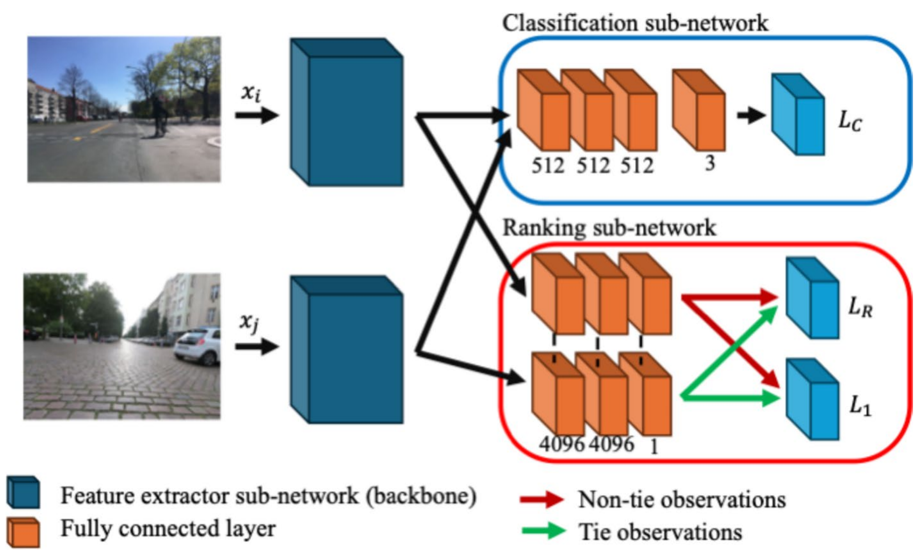

Fig. 7 Architecture of the traffic safety perception prediction model, adapted from Costa et al. (2025a)

dual-task approach allows the model to both classify which image in a pair is safer and assign a continuous perceived safety score to individual images.

The classification sub-network concatenates the feature embeddings of two images and passes them through a fully connected layer. It is trained with a cross-entropy loss (L_{cls}) to predict which image is safer, mirroring the human task of pairwise comparison. The ranking sub-network, on the other hand, maps the feature embedding of a single image to a continuous, unbounded perceived safety score. This sub-network is trained with a ranking loss function (L_{rank}) that enforces a margin γ between the scores of the two images in a pair:

$$L_{rank} = \max(0, s(I_b) - s(I_a) + \gamma) \tag{1}$$

where $s(I_a)$ and $s(I_b)$ are the predicted scores for the safer image I_a and the less safe image I_b, respectively. The overall loss function combines both objectives: $L_{total} = L_{cls} + \lambda L_{rank}$, where λ balances the two tasks. See Costa et al. (2025a) for more details about the safety perception model.

For this study, we use the raw, unnormalized perceived safety scores generated by the ranking sub-network. Costa et al. (2025a) set the margin γ to 0.7 through a hyperparameter search that optimized ranking accuracy on held-out pairwise comparisons. During training, the model incurred a penalty whenever the score difference between a safer and a less safe image fell below 0.7. Because human annotators labeled the training pairs and could reliably distinguish safer from less safe images, a one-unit difference in the resulting score corresponds to a perceptual gap that observers can detect. These scores serve as explanatory variables in our choice models, allowing us to test the direct influence of perceived safety on route choice while controlling for visual street-level features.

Choice model specifications

We estimate four-panel mixed logit discrete choice models. Our utility functions include travel time and traffic lights. Because the experiment by Terra et al. (2025) does not include a monetary cost attribute, willingness-to-pay values cannot be expressed in monetary terms. The omission is also consistent with cycling route choice practice more broadly—for personal bicycle trips, the marginal monetary cost per trip is negligible (no fuel or fares), and cycling route choice studies typically omit cost (Broach et al. 2012; Kuiper 2021). This design choice means our results apply to personal bicycle use and do not extend to shared-bike systems where per-trip costs apply. Traffic lights are included from Model 1 onward as a standard route convenience attribute reflecting intersection delay, not as a safety feature; the safety perception model captures safety independently from the images. These models account for the panel structure in our stated choice data and unobserved taste heterogeneity by allowing key taste parameters to vary randomly across individuals, but not within respondents. For segmentation features, we employ backward stepwise selection to identify the variables contributing most to explanatory power and prioritize model simplicity and interpretability over maximizing fit (e.g., Hou et al. 2019; Madarshahian et al. 2023). The utility functions for the different model specifications are as follows:

(1) Base model:

$$U_{itn} = \beta_{TT,n} \cdot TT_{itn} + \beta_{TL,n} \cdot TL_{itn} + \epsilon_{itn} \tag{2}$$

where U_{itn} is the utility of alternative *i* for decision maker *n* in choice task *t*, TT_{itn} is travel time, TL_{itn} is the number of traffic lights, $\beta_{TT,n}$ and $\beta_{TL,n}$ are individual-specific random taste parameters, and ϵ_{itn} is an IID Extreme Value Type I distributed error term.

(2) Base + Safety model:

$$U_{itn} = \beta_{TT,n} \cdot TT_{itn} + \beta_{TL,n} \cdot TL_{itn} + \beta_{SAFETY,n} \cdot SAFETY_{itn} + \epsilon_{itn} \tag{3}$$

where $SAFETY_{itn}$ represents the perceived traffic safety score, computed using the safety perception model for $IMAGE_{itn}$, for alternative *i* of choice task *t*, and $\beta_{SAFETY,n}$ is an individual-specific random taste random parameter.

(3) Base + Segmentation Model:

$$\begin{aligned} U_{itn} = {} & \beta_{TT,n} \cdot TT_{itn} + \beta_{TL,n} \cdot TL_{itn} \\ & + \sum_k \beta_k \cdot SEG_{k,itn} + \epsilon_{itn} \end{aligned} \tag{4}$$

where $SEG_{k,itn}$ represents the proportion of segmentation feature *k* in the image for alternative *i* presented in choice task *t*, and β_k are (non-random) parameters for the selected segmentation features.

(4) Full Model (Base + Safety + Segmentation):

$$\begin{aligned} U_{itn} = {} & \beta_{TT,n} \cdot TT_{itn} + \beta_{TL,n} \cdot TL_{itn} + \beta_{SAFETY,n} \cdot SAFETY_{itn} \\ & + \sum_k \beta_k \cdot SEG_{k,itn} + \epsilon_{itn} \end{aligned} \tag{5}$$

For the random parameters, we assume negative log-normal distributions for travel time and traffic lights: $\beta_{TT,n} \sim -\exp\left(\mathcal{N}(\mu_{TT}, \sigma^2_{TT})\right)$, $\beta_{TL,n} \sim -\exp\left(\mathcal{N}(\mu_{TL}, \sigma^2_{TL})\right)$, ensuring negative marginal uilities for all *n*, and a normal distribution for safety: $\beta_{SAFETY,n} \sim \mathcal{N}(\mu_{SAFETY}, \sigma^2_{SAFETY})$. The mixed logit probabilities are computed using a Monte Carlo simulation with 1000 Halton draws.

Willingness-to-pay estimation

To compute willingness-to-pay (WTP) measures for safety perception, we estimate separate mixed logit models in WTP space. This approach directly models the trade-offs cyclists are willing to make between safety perception and travel time or traffic lights. We define WTP space utility functions where WTP parameters are estimated with log-normal distributions.

For safety versus travel time:

$$U_{itn} = \mu_{TT} \cdot (TT_{itn} + WTP_{TL,n} \cdot TL_{itn} + WTP_{SAFETY,n} \cdot SAFETY_{itn}) + \sum_k \beta_k \cdot SEG_{k,itn} + \epsilon_{itn} \tag{6}$$

For safety versus traffic lights:

$$U_{itn} = \mu_{TL} \cdot (TL_{itn} + WTP_{TT,n} \cdot TT_{itn} + WTP_{SAFETY,n} \cdot SAFETY_{itn}) + \sum_k \beta_k \cdot SEG_{k,itn} + \epsilon_{itn} \tag{7}$$

where μ_{TT} and μ_{TL} are non-random scale parameters, $WTP_{TL,n} = \exp(\mu_{TL} + \sigma_{TL} \cdot \eta_{TL,n})$, $WTP_{TT,n} = \exp(\mu_{TT} + \sigma_{TT} \cdot \eta_{TT,n})$, $WTP_{SAFETY,n} = -\exp(\mu_{SAFETY} + \sigma_{SAFETY} \cdot \eta_{SAFETY,n})$, and $\eta \sim \mathcal{N}(0,1)$. Mean WTP is computed as $E[WTP] = \exp(\mu + \sigma^2/2)$, representing additional travel time (seconds) or traffic lights cyclists "pay" for a one-unit increase in perceived safety.

Interaction Models

For examining systematic heterogeneity across segments of the cycling population, we estimate mixed logit models with interaction effects. We implement a method where each demographic group has its own random safety parameter, all sharing a common standard deviation. The utility function takes the form:

$$\begin{aligned} U_{itn} = {} & \beta_{TT,n} \cdot TT_{itn} + \beta_{TL,n} \cdot TL_{itn} \\ & + \left[\beta_{SAFETY,ref,n} + \sum_m (\mathbb{I}[DEMO_n = m] \cdot \beta_{SAFETY,m,n}) \right] \cdot SAFETY_{itn} \\ & + \sum_k \beta_k \cdot SEG_{k,itn} + \epsilon_{itn} \end{aligned} \tag{8}$$

where $\beta_{SAFETY,ref,n} = \mu_{SAFETY,ref} + \sigma_{SAFETY,common} \cdot \eta_{ref,n}$ for the reference category, $\beta_{SAFETY,m,n} = \mu_{SAFETY,m} + \sigma_{SAFETY,common} \cdot \eta_{m,n}$ for category m, $\mathbb{I}[\cdot]$ is an indicator function, $DEMO_n$ represents individual n's demographic category, and all groups share the common standard deviation $\sigma_{SAFETY,common}$. The reference categories for each demographic variable are: Age (Young: 18–30 years), Gender (Others), Household Size (Small: 1–2 people), Household Composition (Others), Income (Low: < €2,250), Education (Low), Bills (Easy), Cycling Attitude (Yes: like cycling), Cycling Incident (Severe), Cycling Unsafe Feeling (Sometimes), Cycling Frequency (Occasional: 1–2 days/week), Bike Type (Others), Car Ownership (No car), Transportation Mode (Others), Commuting Days (No commute), Travel Time (No commute), and Trip Purpose (Others).

Results

Main model results

This section addresses our first research question: Does perceived traffic safety affect cyclists' route choices after controlling for visual street-level features extracted through image segmentation? That is, we test whether perceived safety is statistically significant in our models and whether adding perceived safety improves model performance. It does: perceived safety is a positive and significant driver of route choice across specifications, and adding it improves model fit, as we detail below.

Table 4 presents the comprehensive results from benchmarking four different model specifications. The comparison demonstrates clear improvements in model performance when incorporating traffic safety perception and semantic segmentation features. Model performance is primarily evaluated using log-likelihood, which is the standard metric

Table 4 Model benchmark summary

Model type	Model 1	Model 2	Model 3	Model 4
	Base	Base+Safety	Base+Seg	Base+Seg+Safety
Goodness of fit				
Sample size	746	746	746	746
Number of estimated parameters	4	6	15	17
Number of observations	11,190	11,190	11,190	11,190
Log-Likelihood	−7080.88	−6781.60	−6348.73	−6328.27
Rho-squared	0.0871	0.126	0.181	0.184
AIC	14,169.77	13,575.20	12,727.47	12,690.55
BIC	14,188.23	13,602.89	12,796.69	12,769.00
Random Parameter Means (Coefficient, t-value)				
B_TT	−2.21** (−3.13)	−0.64* (−2.43)	0.14 (0.94)	0.18 (1.31)
B_TL	−0.93*** (−4.12)	−0.75** (−3.11)	−0.47* (−2.40)	−0.49** (−2.69)
B_SAFETY_SCORE	–	0.60*** (18.0)	–	0.14*** (3.62)
Segmentation Features				
B_Bike_Lane	–	–	3.99*** (8.93)	3.48*** (7.58)
B_Curb	–	–	9.72*** (4.90)	8.82*** (4.41)
B_Pole	–	–	5.45** (2.95)	5.03** (2.72)
B_Sidewalk	–	–	2.45*** (6.73)	2.09*** (5.57)
B_Sky	–	–	1.85*** (10.6)	1.84*** (10.6)
B_Terrain	–	–	3.86*** (13.5)	3.26*** (10.7)
B_Traffic_Sign_(Front)	–	–	−14.5** (−2.95)	−14.5** (−2.99)
B_Truck	–	–	−4.84*** (−4.77)	−4.89*** (−4.83)
B_Utility_Pole	–	–	−10.9* (−2.35)	−11.8* (−2.54)
B_Vegetation	–	–	2.30*** (16.4)	2.20*** (15.5)
B_Water	–	–	8.22*** (6.67)	7.80*** (6.31)
Random Parameter Standard Deviations				
sigma_TT	3.79*** (5.91)	−2.47*** (−9.54)	1.95*** (13.2)	−1.88*** (−14.0)
sigma_TL	1.93*** (7.41)	−1.79*** (−6.43)	1.61*** (6.73)	−1.60*** (−7.58)
sigma_SAFETY_SCORE	–	0.22*** (3.37)	–	0.20*** (3.36)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

for discrete choice model comparison. Log-likelihood values closer to zero indicate better model fit, while rho-squared provides an additional normalized measure of model fit. Likelihood ratio tests and information criteria (AIC, BIC) support model comparison and performance evaluation.

Our mixed logit models allow for unobserved taste heterogeneity across individuals through random parameters, and the heterogeneity is represented by σ, which shows larger heterogeneity among participants when the value is larger. All models show significant standard deviations for the random parameters, indicating substantial heterogeneity in preferences across cyclists.

Model 1 (Base Model): The base model, including only travel time and traffic lights, achieves modest explanatory power with a log-likelihood of −7080.88 and rho-squared of

0.087. The travel time coefficient is negative and significant (−2.21, $p < 0.01$), as expected, with substantial preference heterogeneity ($\sigma = 3.79$, $p < 0.001$). Similarly, the traffic lights coefficient is negative (−0.93, $p < 0.001$) with significant heterogeneity ($\sigma = 1.93$, $p < 0.001$), thus confirming that cyclists generally prefer routes with fewer traffic lights but with varying preferences across individuals.

Model 2 (Base + Safety): Adding perceived safety substantially improves model performance, with log-likelihood improving to −6781.60 and rho-squared increasing to 0.126. The perceived safety coefficient is positive and highly significant (0.60, $p < 0.001$), confirming that cyclists prefer routes with higher perceived safety. The model also reveals significant heterogeneity in safety preferences ($\sigma = 0.22$, $p < 0.001$), which indicates that while most cyclists value safety, the magnitude of this preference varies considerably across individuals.

Model 3 (Base + Segmentation): Incorporating semantic segmentation features without safety perception achieves substantial improvement with log-likelihood of −6348.73 and rho-squared of 0.181, indicating that detailed environmental characteristics significantly explain route choice. Key negative effects include the presence of trucks (−4.84, $p < 0.001$), utility poles (−10.9, $p < 0.05$), and traffic signs (−14.5, $p < 0.01$), while positive effects emerge from water features (8.22, $p < 0.001$), terrain (3.86, $p < 0.001$), and bike lanes (3.99, $p < 0.001$). The travel time coefficient becomes insignificant (0.14, $p > 0.05$).

Model 4 (Full Model): The full model, combining both safety perception and segmentation features, achieves the highest performance with log-likelihood of −6328.27 and rho-squared of 0.184. The safety coefficient remains positive and highly significant (0.14, $p < 0.001$) even after controlling for detailed environmental features. This provides strong support for the notion that perceived safety is important for cycling route choices. Moreover, we find strong support that safety preferences are heterogeneous across the cycling population: ($\sigma = 0.20$, $p < 0.001$).

To verify that our stepwise feature selection procedure did not lead to overfitting, we performed a train-test split validation using the survey's pre-labeled split: a training set of 606 individuals (9090 observations) and a test set of 140 individuals (2100 observations). As shown in Table 5, the test set achieves a higher rho-squared (0.192) than the training set (0.168), indicating that the model generalizes well to new data. This result confirms that our stepwise selection procedure successfully identified features that contribute to model performance without overfitting to the training data. We also performed a comprehensive examination of interaction effects across 17 demographic variables to understand preference heterogeneity. While this extensive testing could potentially increase the risk of finding statistically significant results by chance, we note that many of the interaction effects found in our study align with prior research and theoretical expectations.

The willingness-to-pay estimates reveal cyclists have strong preferences for safer cycling environments. From our WTP space models (Table 6), cyclists are willing to accept 64 additional seconds of travel time and 0.59 additional traffic lights for each unit increase in perceived safety, even after accounting for a rich set of segmentation features. These

Table 5 Train versus Test Model Performance Comparison

	Training	Test
Number of observations	9090	2100
Number of individuals	606	140
LL	−5160.21	−1176.03
ρ^2	0.168	0.192

estimates are derived from the mixed logit models in WTP space with log-normal WTP distributions. Given that safety scores range from approximately −2 (very unsafe) to +2 (very safe), cyclists would accept substantial increases in travel time and traffic lights to experience improvements in cycling conditions from the worst to best perceived safety conditions.

It is important to note that these WTP models show a slightly worse fit than Model 4 (the full utility space model). This is because the WTP models have one fewer random parameter. In Model 4, we estimated travel time as a random parameter, which contributes substantially to model fit. When we estimate a mixed logit model in utility space with travel time as a fixed coefficient, we observe a similar fit to the WTP models. We intentionally keep the scale parameter (either travel time or traffic light) as a fixed parameter in WTP models rather than a random parameter. While allowing the scale parameter to be random would improve model fit, it would confound scale heterogeneity with taste heterogeneity (Hess and Rose 2012), undermining our goal of obtaining meaningful WTP estimates.

The qualitative analysis of model predictions provides additional insights into how safety perception improves utility estimation. Figure 8 shows images where the model incorporating safety perception makes different utility predictions compared to the model using only segmentation features. The plots on the right side show the percentiles of the images in terms of estimated utility among all the images. In the upper panel, images where the model with perceived safety estimated a higher utility than the model without perceived safety seem to have good visibility of the roads and appear organized with low speed limits for vehicles, even if cyclists have to share the road with vehicles. In the lower panel, images where the model with perceived safety estimated lower utility than the model without perceived safety seem to have lower visibility due to lighting and objects on the road (e.g., parked cars and bridges), and appear unclear regarding where cyclists are protected from seemingly high-speed roads.

Similarly, Fig. 9 shows pairs of images in which the model with perceived safety predicted the choice of routes by the participants correctly, while the model without perceived safety failed. The bar plots show the probability of the choice predicted by both models. In these images, we can also make similar observations. The model with perceived safety seems to get hints from contextual features that might not appear in the segmentation features, such as the roads with dead ends.

Relationships among utility, perceived safety, and segmentation features (RQ1, RQ2)

Figure 10 shows relationships between route utility from Model 4, perceived safety, and key segmentation features. Positive correlations exist between perceived safety and infrastructure/natural elements (terrain, bike lane, vegetation), while negative correlations emerge with traffic-related features (car, road). Cyclist preferences align with environments that

Table 6 Mixed Logit WTP Space Model Results

Trade-off	Mean WTP	μ	σ	Log-Likelihood
Safety versus Travel Time	1.07 min	−3.090	−1.308	−6785.60
Safety versus Traffic Lights	0.59 lights	−2.257	−1.125	−6485.58

Mixed logit WTP space models with log-normal WTP distribution. Mean WTP = $\exp(\mu + \sigma^2/2)$. Includes significant segmentation features

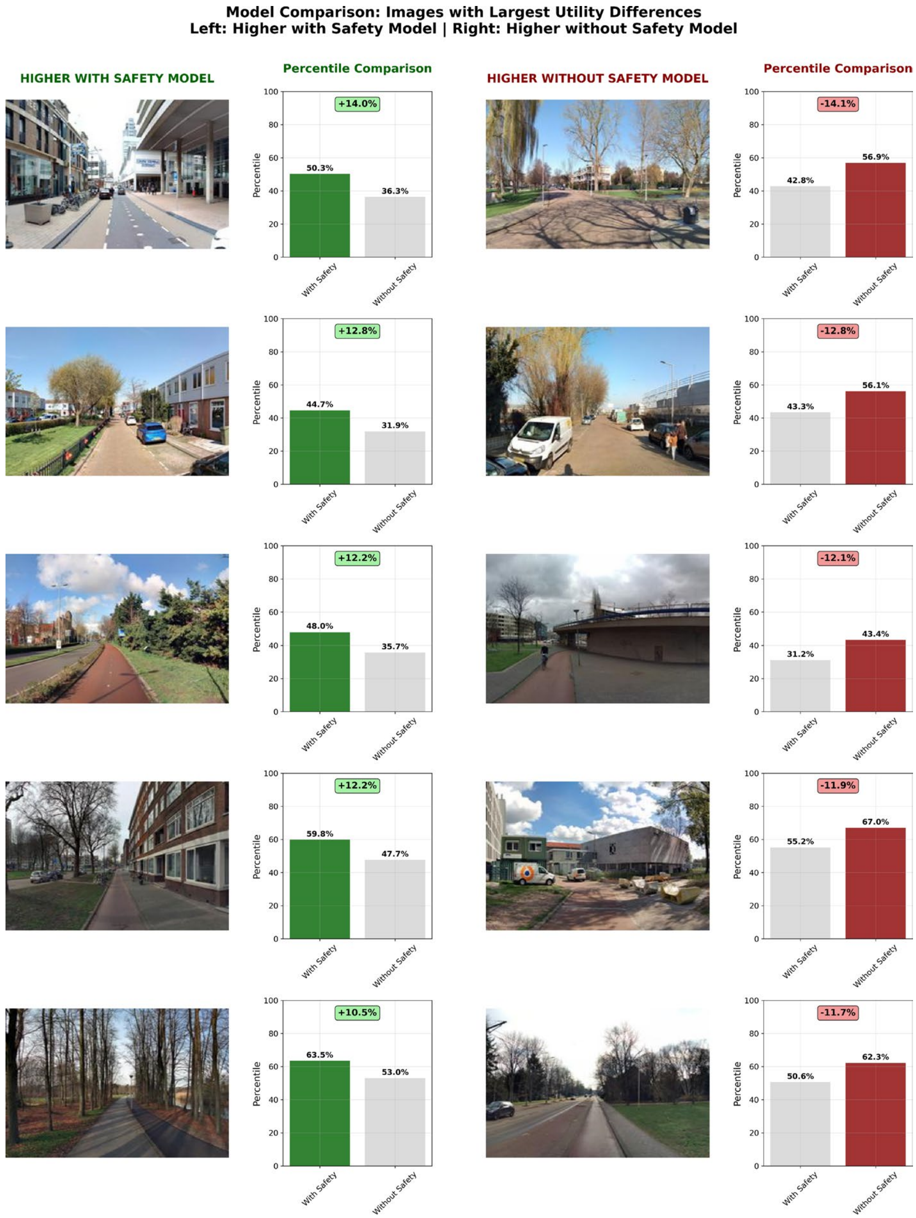

Fig. 8 Qualitative comparison of utility estimation: images where the model with perceived safety estimated a different utility than the model without perceived safety

provide good visibility while maintaining greenery, suggesting a preference for open terrain over dense vegetation that might obstruct sightlines.

Some segmentation features are correlated with each other (e.g., vegetation and terrain tend to co-occur; car and road proportions are positively related). Such correlations are common when predictors describe the same scene and, as Fig. 10 shows, none of the retained

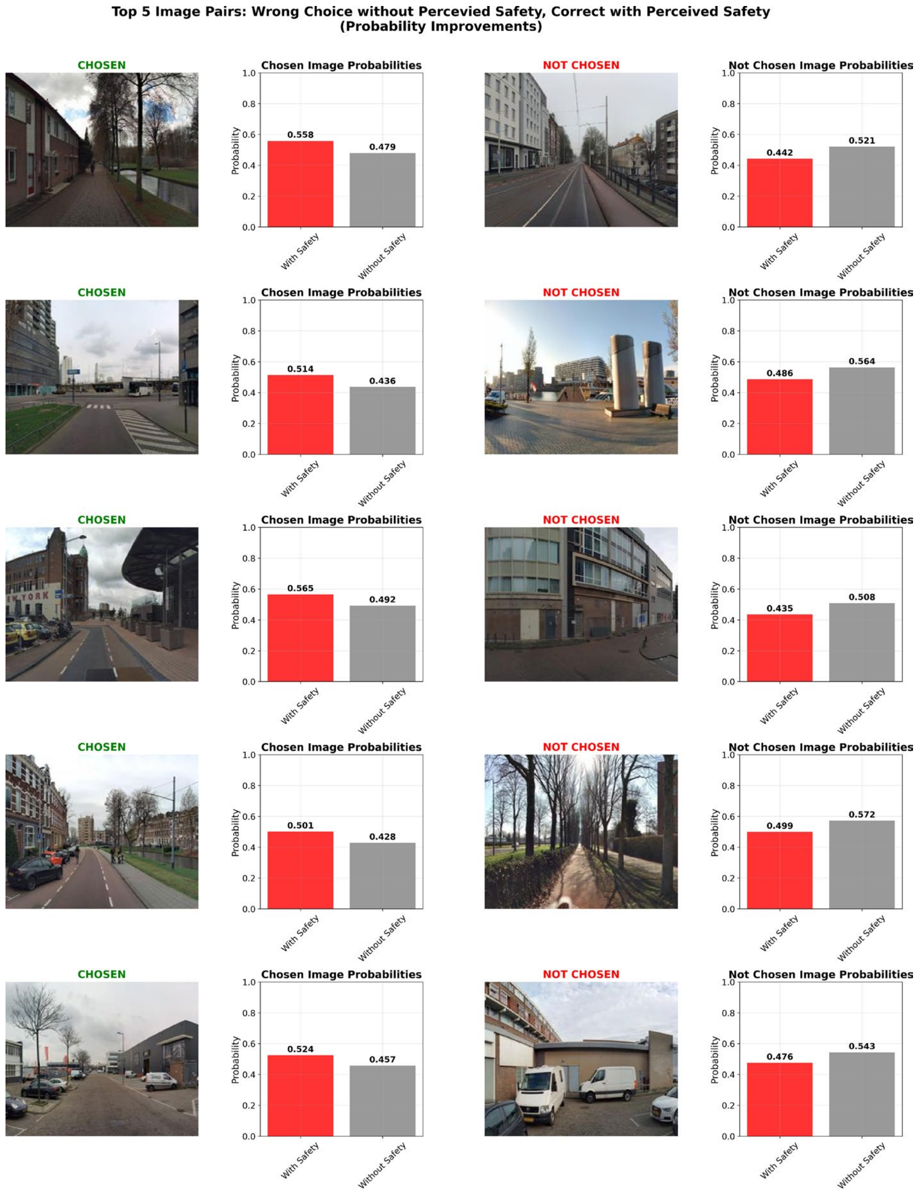

Fig. 9 Qualitative comparison of route choice prediction: images where the model with perceived safety predicted the choice of routes by the participants correctly, while the model without perceived safety failed

features exceeds the conventional threshold ($|r| > 0.9$) at which collinearity impedes parameter identification. Moderate correlations inflate standard errors but do not bias coefficient estimates, and attributes are not required to be independent in a discrete choice model. The stepwise selection we use on top of this is motivated by model parsimony rather than by

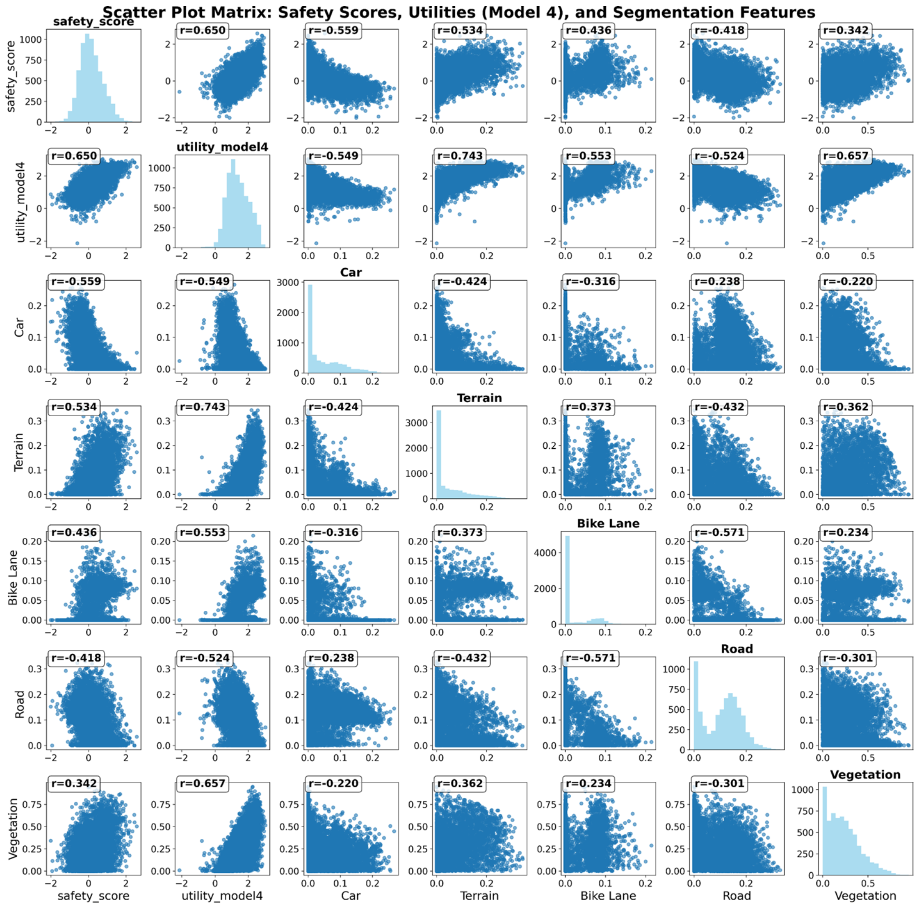

Fig. 10 Scatter plot matrix showing relationships between utility, perceived safety, and key segmentation features

a multicollinearity concern, and the train-test validation (Table 5) confirms that the retained features generalize to new data.

Heterogeneous effects of traffic safety perception (RQ3)

This subsection shows the results that answer our third research question: To what extent do preferences for perceived traffic safety vary across demographic groups, and which groups attach relatively greater or lower importance to safer cycling routes? To examine how safety preferences vary across different segments of the cycling population, we estimated 17 interaction models between perceived safety and demographic variables covering individual characteristics (age, gender), household factors (size, composition), socioeconomic status (income, education, bills), cycling experience (attitude, incident history, unsafe feelings, frequency, bike type), transportation patterns (car ownership, primary mode), and trip characteristics (commuting days, travel time, trip purpose). These models help identify which groups would benefit most from safety-focused infrastructure investments and reveal important heterogeneity in cycling preferences that can inform targeted policy interventions.

The interaction effects are presented in Tables 7, 8, and 9, which show the safety-related parameters from all 17 interaction models. Since the interaction models are generalizations of Model 4, we assess their improved performance using a Likelihood Ratio Test. The results indicate that only three interaction models provide statistically significant improvements in fit over Model 4 at $\alpha = 0.05$: Age (LRS = 6.12, critical $\chi^2_{(2,0.05)}$ = 5.99), Cycling Attitude (LRS = 6.71, critical $\chi^2_{(1,0.05)}$ = 3.84), and Trip Purpose (LRS = 15.00, critical $\chi^2_{(3,0.05)}$ = 7.81). While several other models show significant individual interaction parameters, these three models demonstrate both statistically significant interaction effects and meaningful improvements in overall model performance. Below, we discuss the estimation results with

Table 7 Safety-demographics interaction models summary (Part 1: Demographics and Socioeconomic)

	Age	Gender	HH size	HH comp.	Income
Model Fit Statistics					
Sample size	746	746	746	746	746
Parameters	19	19	19	19	19
Log-Likelihood	−6325.22	−6327.81	−6326.26	−6327.58	−6328.22
LRT Stat	6.12*	0.94	4.03	1.40	0.11
Safety interaction parameters: ref (t-stat vs. 0), others (z-stat vs. ref)					
Young (ref)	0.02 (0.27)	–	–	–	–
Middle	0.10 (0.68)	–	–	–	–
Senior	0.21 (1.51)	–	–	–	–
Others (ref)	–	0.62 (0.57)	–	0.03 (0.33)	–
Female	–	−0.51 (−0.73)	–	–	–
Male	–	−0.47 (−0.70)	–	–	–
Alone	–	–	–	0.17 (0.85)	–
Couple	–	–	–	0.10 (0.44)	–
Small (ref)	–	–	0.20*** (4.23)	–	–
Medium	–	–	−0.18*** (−4.58)	–	–
Large	–	–	0.06 (−1.02)	–	–
Low (ref)	–	–	–	–	0.17* (2.16)
Medium	–	–	–	–	0.02 (−1.22)
High	–	–	–	–	−0.09* (−2.18)

LRT versus Model 4: Age (Yes*), Gender (No), HH Size (No), HH Comp (No), Income (No)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 8 Safety-demographics interaction models summary (Part 2: Socioeconomic and Cycling Experience)

	Educ.	Bills	Attitude	Incident	Unsafe	Freq.
Model Fit Statistics						
Sample	746	746	746	746	746	746
Params	19	19	18	19	19	18
LL	−6328.12	−6328.03	−6324.92	−6327.97	−6326.88	−6327.94
LRT	0.31	0.49	6.71*	0.60	2.79	0.66
Safety parameters: ref (t vs. 0), others (z vs. ref)						
Low (ref)	0.18 (1.82)	–	–	–	–	–
Med	−0.01 (−1.27)	–	–	–	–	–
High	−0.07 (−1.74)	–	–	–	–	–
Easy (ref)	–	0.19*** (4.03)	–	–	–	–
Reasonable	–	−0.10*** (−3.45)	–	–	–	–
Difficult	–	−0.09* (−2.14)	–	–	–	–
Yes (ref)	–	–	0.19*** (4.62)	–	–	–
No	–	–	−0.27*** (−4.55)	–	–	–
Severe (ref)	–	–	–	0.13 (1.00)	–	–
Mild	–	–	–	−0.01 (−0.70)	–	–
No	–	–	–	0.01 (−0.66)	–	–
Sometimes (ref)	–	–	–	–	0.12 (1.61)	–
Eve/Night	–	–	–	–	−0.04 (−1.36)	–
No	–	–	–	–	0.10 (−0.19)	–
Occasional (ref)	–	–	–	–	–	0.16** (2.77)
Regular	–	–	–	–	–	−0.03* (−2.08)

LRT versus Model 4: Educ (No), Bills (No), Attitude (Yes*), Incident (No), Unsafe (No), Freq (No)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

emphasis on these three better-performing models, followed by brief summaries of notable patterns in the remaining models.

Age: Age emerges as a key determinant of safety preferences, with the interaction model showing statistically significant improvement over the base model. While the reference group of young cyclists (18–30 years) shows no significant safety preference (0.02, p >0.05), older cyclists demonstrate increasing value for safety. Middle-aged cyclists (31–60) show a positive coefficient (0.10), though this difference from young cyclists is not statistically significant (z = 0.68, p >0.05). Senior cyclists aged 61 and above show a stronger preference for safety (0.21), and this difference from young cyclists approaches significance (z = 1.51, p >0.10). One tentative interpretation, which our cross-sectional design cannot identify, is that older cyclists may be more risk-averse or have reduced physical confidence when navigating potentially unsafe cycling environments. These findings have important implications for designing inclusive cycling infrastructure that

Table 9 Safety-demographics interaction models summary (Part 3: Cycling Type, Transportation and Trip)

	BikeType	Car	Trans	CommDays	TravelTime	TripPurp
Model Fit Statistics						
Sample	746	746	746	746	746	746
Params	20	19	21	19	19	20
LL	−6326.03	−6328.21	−6327.60	−6327.24	−6328.67	−6320.77
LRT	4.50	0.13	1.36	2.06	−0.79	15.00*
Safety parameters: ref (t vs. 0), others (z vs. ref)						
Others (ref)	−0.27 (−0.84)	–	0.24 (0.47)	–	–	−0.16 (−1.02)
E-bike	0.41 (1.51)	–	–	–	–	–
Racing	0.26 (1.14)	–	–	–	–	–
Regular	0.43 (1.55)	–	–	–	–	–
Bike	–	–	−0.12 (−0.51)	–	–	–
Car	–	–	−0.10 (−0.48)	–	–	–
Public	–	–	−0.06 (−0.41)	–	–	–
Walk	–	–	−0.01 (−0.35)	–	–	–
No car (ref)	–	0.24** (2.66)	–	–	–	–
One car	–	−0.13** (−2.78)	–	–	–	–
Multi cars	–	−0.06* (−2.07)	–	–	–	–
No commute (ref)	–	–	–	0.20*** (3.40)	0.22*** (3.68)	–
Part-time	–	–	–	−0.18*** (−3.71)	–	–
Full-time	–	–	–	−0.02* (−2.36)	–	–
Short	–	–	–	–	−0.11** (−3.19)	–
Long	–	–	–	–	−0.12*** (−3.58)	–
Commute	–	–	–	–	–	0.28 (1.91)
Errands	–	–	–	–	–	0.25 (1.79)
Recreational	–	–	–	–	–	0.49** (2.79)

LRT versus Model 4: BikeType (No), Car (No), Trans (No), CommDays (No), TravelTime (No), TripPurp (Yes*)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

can attract and retain older adults, who represent a growing demographic with significant untapped cycling potential.

Cycling attitudes: Cycling attitudes strongly predict safety preferences, with the interaction model showing statistically significant improvement over the base model. Cyclists who like cycling show significantly positive safety preferences (0.18, p <0.001), while those who report not liking cycling place significantly less value on safety (−0.27, z = −4.55, p<0.001

relative to those who like cycling). This substantial difference suggests that cycling attitudes fundamentally shape how individuals weigh safety in their route choices. The counterintuitive finding for those who dislike cycling suggests they may prioritize other factors (such as speed or convenience) over safety, or may have already adapted to unsafe conditions. From a policy perspective, this indicates that safety improvements alone may be insufficient to convert reluctant cyclists who have fundamental negative attitudes toward cycling.

Trip purpose: Trip purpose reveals the most substantial dimension of preference heterogeneity, with the interaction model showing the greatest statistical improvement over the base model (LRS = 15.00). Using “Others” (including exercise, picking up/dropping off, and other miscellaneous purposes) as the reference category, which shows no significant safety preference (−0.16, p >0.05), we observe meaningful differences across trip purposes. Recreational cyclists place substantially higher value on safety (0.49, z = 2.79, p <0.01) compared to the reference group, while commuting (0.28, z = 1.91) and errands (0.25, z = 1.79) show positive but not statistically significant differences. This pattern makes intuitive sense: recreational cycling is voluntary and discretionary, allowing cyclists more freedom to choose safer (though potentially longer) routes, while trips with other purposes may face greater time or routing constraints.

Other models with notable interaction parameters: While 14 other interaction models did not show statistically significant improvement over the base model based on LRT, several reveal interesting patterns in their parameter estimates. Household size (LRS = 4.03) shows small households (1–2 people, 0.20, p <0.001) with higher safety preferences than medium households (−0.18, z = −4.58, p < 0.001). Bike type (LRS = 4.50) shows regular bikes (0.43), e-bikes (0.41), and racing bikes (0.26) with higher safety preferences versus other types (−0.27). Bills (LRS = 0.49) shows those who find bills easy (0.19, p < 0.001) valuing safety more than those finding it reasonable (−0.10, z = −3.45, p <0.001) or difficult (−0.09, z = −2.14, p <0.05). Travel time (LRS = −0.79) and commuting days (LRS = 2.06) show those without commutes valuing safety more. Income (LRS = 0.11), education (LRS = 0.31), car ownership (LRS = 0.13), cycling frequency (LRS = 0.66), cycling incident history (LRS = 0.60), unsafe feelings (LRS = 2.79), transportation mode (LRS = 1.36), gender (LRS = 0.94), and household composition (LRS = 1.40) show no significant model improvement, though some individual parameters are significant.

Discussion

Does perceived traffic safety affect cyclists’ route choices after controlling for visual street-level features?

Perceived safety remains a positive and significant predictor of route choice even after controlling for a rich set of visual street-level features extracted through semantic segmentation (RQ1). Human perception processes visual information in ways that go beyond the presence or absence of individual scene components. Semantic segmentation captures the proportion of bike lanes, vegetation, or vehicles in an image, but not subtler cues like pavement condition, visual clutter, sight lines, or the overall sense of order that shape safety judgments. The residual explanatory power of safety scores after controlling for these features confirms that perceived safety captures a dimension of route quality that component-level analysis misses. Two non-exclusive mechanisms can produce this residual. First, segmentation is an incomplete inventory. It omits measurable but uncoded cues such as pavement defects, lighting, parked vehicles, and the legibility of where a cyclist belongs, so the safety score may stand in for omitted variables that a richer feature set would eventually capture. Second, and more fundamentally, safety perception appears to be configural rather than additive. Riders read a scene as a whole, in which the same element reassures in one spatial arrangement and threatens in another, and the perceived-safety score, trained on holistic human judgments, encodes interactions that a sum of independent component proportions cannot. Cross-sectional data cannot fully separate the two mechanisms, but both imply that scoring the cycling environment as experienced, rather than decomposing it into parts, better reflects how cyclists evaluate routes.

The finding contrasts with Huber et al. (2024), who found that perceived safety (measured through survey-reported incidents) did not significantly influence route choices in a revealed preference study. The discrepancy may reflect either methodological differences (their RP approach captures behavior constrained by the existing network, while our SP design presents controlled trade-offs that can surface latent preferences) or measurement differences (self-reported critical incidents and image-derived perception scores may capture different constructs). The relative reliability of the two approaches for recovering safety preferences remains an open question for further research.

The willingness-to-pay estimates put a concrete value on safety perception: cyclists accept about 64 additional seconds of travel time (roughly 11% of a 10-min trip) per unit increase in perceived safety. For context, Broach et al. (2012) found that cyclists in Portland would travel up to 26% farther to use an off-street bike path, and Lu et al. (2025) reported that visual features of the built environment alter perceived route distance by up to 60%. Our WTP estimate complements these distance-based measures by expressing the trade-off in time units, which are more directly useful for transport appraisal.

The WTP estimate also fits the worthwhile travel time framework reviewed by Cornet et al. (2022), which argues that travel time is not purely a cost when conditions make the journey enjoyable, productive, or comfortable. The WTP magnitude alone does not identify which side of the trade-off carries more weight: we cannot tell from the ratio whether cyclists place a high utility on safety or a low disutility on travel time. Either way, the accepted additional travel time ($\approx$ 64 s per safety unit) makes the time spent cycling on a safer route feel worthwhile rather than wasted, so the WTP estimate can be interpreted not just as a disutility trade-off but as the value of travel conditions that transform time spent cycling from stressful to acceptable.

To what extent do preferences for perceived traffic safety vary across demographic groups?

Safety preferences vary across the cycling population (RQ3). Of the 17 demographic interaction models we estimated, three produced statistically significant improvements over the base model: age, cycling attitude, and trip purpose. Several other dimensions showed individually significant parameters without improving the overall model fit. We discuss the key patterns below, focusing on interpretation and comparison with prior work rather than restating numerical results already reported in Results section.

Age. Older cyclists place greater weight on safety when choosing routes, consistent with Uijtdewilligen et al. (2024), who reported that older cyclists are more sensitive to crowding and route safety attributes in a stated-preference experiment, and with Rossetti and Daziano (2023), whose latent-class segmentation by self-assessed health (which co-varies with age) recovered classes that differ in their willingness to trade safety against time. de Freitas and Axhausen (2025) and Keppner et al. (2023) similarly find that older adults feel less safe cycling and are more supportive of safety-oriented infrastructure investment, and Sener et al. (2009) pointed to age as a primary axis of preference heterogeneity in an early demographic segmentation of cycling preferences. Kuiper (2021) reports a similar age effect in a Dutch RP study. While this age gradient is well established in the literature, our contribution is to quantify it within a WTP framework: the difference in safety valuation between younger and senior cyclists translates directly into how much additional travel time each group would accept for a safer route, information that planners can use when prioritizing infrastructure upgrades in areas with older populations.

Cycling attitudes. Cyclists who report liking cycling place more weight on safety than those who do not. At first glance, this seems counterintuitive: one might expect people who dislike cycling to be more sensitive to safety concerns. However, the relationship between attitudes and safety valuation is likely bidirectional (Berghöfer and Vollrath 2023). People who enjoy cycling may have developed a finer appreciation for route quality differences, including safety, through accumulated experience. Conversely, those who dislike cycling may be "forced" cyclists (e.g., because they lack a car) who prioritize speed and convenience over safety, or they may have psychologically adapted to unsafe conditions. We cannot establish causality with our cross-sectional design, and cycling experience (years of cycling) was not collected as a separate variable. Cycling frequency and cycling attitude likely serve as proxies for experience, but future longitudinal designs would be needed to disentangle cause and effect. All respondents are Dutch, and cycling is a mass-practiced everyday skill typically acquired in childhood in the Netherlands (Pucher and Buehler 2008), so the range of cycling proficiency within our sample is probably compressed relative to countries with lower cycling penetration.

Trip purpose. Recreational cyclists value safety far more than commuters or errand-runners, and this interaction model showed the greatest statistical improvement. Broach et al. (2012) found a similar split: commuters are more distance-sensitive, while non-commuters are more willing to detour for better infrastructure. Grüner and Vollrath (2021) similarly identified two commuter cyclist types, "foremost fast" and "safe and fast," and found that efficiency dominated actual route choices for both groups. Recreational trips are discretionary: cyclists can freely choose safer (if longer) routes, and the trip itself is part of the experience rather than a means to reach a destination. Commuters face tighter time budgets and are more likely to accept less pleasant routes to minimize delay.

Household size and financial stability. Small households (1–2 people) show stronger safety preferences than larger ones, and respondents who report finding their bills easy to pay also show higher safety valuations than those who do not. We offer two tentative mechanisms, which our cross-sectional design cannot confirm. People living alone or as a couple might have more flexible schedules with fewer childcare or school-run obligations, leaving room to choose longer but safer routes. The bills variable, which captures self-reported ease of paying household bills, complements income brackets as a grounded measure of financial stability; financial security could plausibly reduce the pressure to minimize travel time, freeing individuals to prioritize route quality. Both mechanisms are consistent with the data, but neither is identified by it.

Dimensions without significant effects. Gender, education, cycling incident history, and cycling unsafe feelings did not yield significant interaction effects, suggesting safety preferences are relatively uniform across these dimensions. The absence of a gender effect is worth noting, as Uijtdewilligen et al. (2024) found gender differences in route preference. One explanation is that our study captures only the safety component of route choice; gender differences may manifest through other attributes (e.g., crowding, lighting) not isolated in our interaction models.

These heterogeneous effects could reflect two mechanisms: certain groups may perceive environments as less safe (while weighing safety equally), or they may weigh safety more heavily regardless of perception. Our design, which uses a common CV-derived safety score for all respondents, does not disentangle these channels. Regardless of the mechanism, the policy implication is the same: groups showing stronger safety effects stand to benefit most from safety-oriented infrastructure improvements.

Limitations

This study has several limitations. First, our street-level imagery represents static snapshots and may not capture how safety perceptions vary with weather, lighting, or traffic conditions throughout the day. Each route was also represented by a single image, which may not reflect the full experience of longer routes where conditions change along the way. Seasonal changes in vegetation, infrastructure maintenance, or temporary construction are similarly absent, which may weaken external validity for real-world scenarios where perceptions shift dynamically.

Second, the safety perception model was trained on a younger Danish and German population, which may introduce systematic differences in how Dutch participants perceived the images. Interpretations of safety cues, familiarity with infrastructure types, and cultural associations with cycling environments may differ across populations. Our model validation showed reasonable patterns across Rotterdam road types, but systematic bias remains possible. Analysts should exercise caution when applying these scores to populations that differ from the training data, and perception models may require local calibration.

Third, omitted variable bias may be present. Safety scores likely correlate with other perceptual factors not explicitly measured, such as beauty, cleanliness, or perceived neighborhood wealth. Our segmentation features capture the most prominent visual elements but may miss cues that influence both safety perception and route choice. Because urban environmental qualities are correlated, the estimated safety effect could partly reflect unmeasured factors, risking overestimation.

Fourth, stepwise feature selection and testing 17 demographic interactions sequentially raises the possibility that some significant results occurred by chance. A train-test split validation confirmed that the model generalizes to unseen data, and many interaction effects align with prior research, but replication in independent samples would strengthen confidence in the demographic-specific findings.

Fifth, transferability to contexts outside the Netherlands and Germany deserves attention. Both the safety perception model (trained in Berlin) and the choice experiment (conducted in the Netherlands) reflect countries with developed cycling infrastructure, though the two settings differ in modal share and infrastructure style: the Netherlands has much higher cycling penetration and more fully separated networks, while Germany leans more on mixed-traffic provision (Pucher and Buehler 2008). In cities with less developed cycling networks, safety perceptions may be shaped by different cues, and the WTP estimates could differ. That said, some safety-related preferences appear robust across cultures: separation from motor traffic is valued in studies from Portland (Broach et al. 2012), the Netherlands (Uijtdewilligen et al. 2024), and Germany (Berghöfer and Vollrath 2022). The methodological framework itself (SP + CV-derived safety scores + WTP estimation) is transferable, even if the specific model coefficients require local calibration. In the Rotterdam context specifically, the finding that main roads receive higher safety scores than residential roads (Fig. 3) reflects a local characteristic: major roads in Rotterdam typically have well-separated cycle paths, whereas residential streets often lack dedicated cycling infrastructure, providing a mixed traffic experience.

Conclusion

This study contributes to the literature on cycling route choice by integrating a computer vision-derived safety perception score into a discrete choice modeling framework estimated on stated preference data, enabling willingness-to-pay estimation and systematic heterogeneity analysis. To our knowledge, no prior study has combined these elements for cycling route choice. Cyclists value safety improvements equivalent to 64 s of additional travel time per unit increase in perceived safety, providing a quantitative basis for infrastructure investment decisions.

The heterogeneous effects analysis reveals variation in safety preferences across demographic groups. Senior cyclists aged 61 and above, lower-income individuals, occasional cyclists, those from small households, those without cars, non-commuters, and recreational cyclists all place a higher value on safety improvements, suggesting that safety interventions may be particularly effective for promoting inclusive cycling participation among these groups.

Beyond behavioral insights, the framework enables three practical applications for infrastructure planning. First, planners can apply the safety perception model to score existing cycling networks from street-level imagery, identifying low-scoring segments for targeted improvement. Second, the WTP estimate (64 s per safety unit) can be used in cost-benefit analysis: for a route serving 1,000 daily cyclists, a one-unit safety improvement would generate the equivalent of 17.8 person-hours per day in time-savings benefits. Third, as image generation techniques advance (Anokhin et al. 2020), synthetic images of proposed road designs could be scored before construction, enabling scenario planning for infrastructure that does not yet exist; this extends the CV-DCM framework of Terra et al. (2025) by adding an explicit perceived-safety evaluation. The segmentation coefficients further reveal which visual elements drive route utility (e.g., vegetation, bike lanes, absence of trucks), giving planners specific design levers. Together, these tools move perceived safety from a descriptive concept to a quantifiable input for evidence-based cycling infrastructure investment.

Future research should pursue several promising directions to extend and improve upon our findings. First, future studies could test how safety perceptions vary across different temporal contexts and how these variations affect route choice decisions to address the limitation of static image representation by modifying the time of the day and seasons of the input images with image generation models such as the one by Anokhin et al. (2020). Second, the integration of physiological and virtual reality methods could provide complementary insights into the cognitive and emotional processes underlying safety perception and route choice decisions. Bogacz et al. (2021) used an electroencephalography device and a virtual reality device to examine the cyclists' perceptions and behaviors—future studies can also leverage such devices to test the effects of different urban designs on route choices, which can offer a deeper understanding of the mechanisms through which environmental features influence cyclist behavior.

Third, future work should explore the transferability of our findings across different urban contexts and cultural settings. Just as evidenced by Quintana et al. (2025), the relationship between environmental characteristics and cycling safety perceptions may vary across different countries and cycling cultures, and comparative studies could identify universal versus context-specific factors influencing cycling route choice. Cross-cultural validation of safety perception models could improve their transferability across different urban contexts and demographic groups.

Fourth, future research could comprehensively quantify different perceptual indicators to help disentangle correlations between safety, beauty, cleanliness, and perceived neighborhood wealth, providing a more complete understanding of environmental factors influencing route choice. This would help address potential omitted variable bias and isolate the specific contribution of safety perception to route choice behavior.

This study demonstrates the potential for combining advanced computer vision techniques with established discrete choice modeling frameworks to address complex urban planning challenges. As cities worldwide seek to promote sustainable transportation and improve urban livability, our approach provides a scalable and interpretable method for understanding and improving cycling infrastructure to serve diverse urban populations.

Appendix A. Additional Tables

Additional detailed model results, including parameter estimates for all segmentation features and robustness checks, are available from the authors upon request (Table 10).

Table 10 Example of demographics data collected in the survey

Variable	Respondent 1	Respondent 2	Respondent 3	Respondent 4
Age	31–45 years	46–60 years	18–30 years	18–30 years
Gender	Female	Female	Female	Female
Household Composition	One adult with children	Two or more adults (not couple)	Live alone	Two or more adults (not couple)
Household Size	4 people	2 people	1 person	3 people
Education	Lower secondary	M.Sc	Intermediate vocational	Intermediate vocational
Income	€1,251-€1,700	€1,701-€2,250	€2251-€3650	€3651-€7,000
Bills	Reasonable	Reasonable	Reasonable	Reasonable
Transportation Mode	Bike	Bike	Car	Bike
# of Cars Owned	No cars	1 car	1 car	2 cars
Travel Time	No commute	10–20 min	10–20 min	20–30 min
# of Commuting Days	No commute	5+ days/week	5+ days/week	2 days/week
Cycling Frequency	5+ days/week	5+ days/week	3 days/week	4 days/week
Cycling Accident	No	No	No	Yes, mild
Cycling Attitude	Yes	Yes	Yes	No
Cycling Unsafe	Yes, sometimes	Yes, sometimes	Yes, evening/night	Yes, evening/night
Bike Type	Regular bike	Regular bike	Regular bike	Regular bike
Trip Purpose	Commuting	Commuting	Errands	Commuting
Choice task 1	1	2	2	2
..	..	..	..	..
Choice task 15	2	1	2	2

Acknowledgements We thank all the study participants who took part in the stated choice experiment. We are grateful to the City of Rotterdam for providing the street-level images that formed the foundation of this research.

Author Contribution Koichi Ito: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. Roos Terra: Conceptualization, Data curation, Investigation, Methodology, Writing – review & editing. Miguel Costa: Methodology, Software, Writing – review & editing. Felix Wilhelm Siebert: Methodology, Writing – review & editing.

Francisco Garrido-Valenzuela: Methodology, Writing – review & editing. Carlos Lima Azevedo: Methodology, Writing – review & editing. Filip Biljecki: Conceptualization, Funding acquisition, Project administration. Sander van Cranenburgh: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.

Funding This research was funded by the Singapore International Graduate Award (SINGA) scholarship provided by the Agency for Science, Technology, and Research (A*STAR) and by the TU Delft AI Lab Programme. This research is part of the project Large-scale 3D Geospatial Data for Urban Analytics, which is supported by the National University of Singapore under the Start Up Grant R-295-000-171-133.

Data availability The street-level images and stated-preference choice data used in this study are owned by Terra et al. (2025) and may be requested from those authors. The perceived-safety model is built on the Berlin perception data of Costa et al. (2025a).

Code availability The replication code is available at https://github.com/koito19960406/cycling_safety_perception.

Declarations

Conflict of interest The authors have no conflict of interest to declare that are relevant to the content of this article.

Ethics approval Not applicable.

Generative AI and AI-assisted technologies in the writing process During the preparation of this work, the author(s) used Claude Sonnet 4 to format tables and proofread. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

References

Anokhin, I., Solovev, P., Korzhenkov, D., Kharlamov, A., Khakhulin, T., Silvestrov, A., Nikolenko, S., Lempitsky, V., Sterkin, G.: High-Resolution Daytime Translation Without Domain Labels. (2020). arXiv:2003.08791 [cs]

Berghöfer, F.L., Vollrath, M.: Cyclists' perception of cycling infrastructure - a repertory Grid approach. Transport. Res. F: Traffic Psychol. Behav. **87**, 249–263 (2022). https://doi.org/10.1016/j.trf.2022.04.012

Berghöfer, F.L., Vollrath, M.: Motivational and deterrent effects of route attributes in cyclists' route choice. Transport. Res. F: Traffic Psychol. Behav. **95**, 343–354 (2023). https://doi.org/10.1016/j.trf.2023.05.004

Biljecki, F., Ito, K.: Street view imagery in urban analytics and GIS: a review. Landsc. Urban Plan. **215**, 104217 (2021). https://doi.org/10.1016/j.landurbplan.2021.104217. (https://www.sciencedirect.com/science/article/pii/S0169204621001808)

Blondiau, T., van Zeebroeck, B., Haubold, H.: Economic benefits of increased cycling. Transp. Res. Procedia **14**, 2306–2313 (2016). https://doi.org/10.1016/j.trpro.2016.05.247. (https://www.sciencedirect.com/science/article/pii/S2352146516302538)

Bogacz, M., Hess, S., Calastri, C., Choudhury, C.F., Mushtaq, F., Awais, M., Nazemi, M., van Eggermond, M.A.B., Erath, A.: Modelling risk perception using a dynamic hybrid choice model and brain-imaging data: an application to virtual reality cycling. Transp. Res. C: Emerg. Technol. **133**, 103435 (2021). https://doi.org/10.1016/j.trc.2021.103435. (https://www.sciencedirect.com/science/article/pii/S0968090X21004277)

Bovy, P.H.L., Bradley, M.A.: Route choice analyzed with stated-preference approaches. Transportation Research Record (1985). https://trid.trb.org/View/272229. iSBN: 9780309039536

Brand, C., Dons, E., Anaya-Boig, E., Avila-Palencia, I., Clark, A., de Nazelle, A., Gascon, M., Gaupp-Berghausen, M., Gerike, R., Götschi, T., Iacorossi, F., Kahlmeier, S., Laeremans, M., Nieuwenhuijsen, M.J., Pablo Orjuela, J., Racioppi, F., Raser, E., Rojas-Rueda, D., Standaert, A., Stigell, E., Sulikova, S., Wegener, S., Int Panis, L.: The climate change mitigation effects of daily active travel in cities. Transp. Res. Part D: Transp. Environ. **93**, 102764 (2021). https://doi.org/10.1016/j.trd.2021.102764. (https://www.sciencedirect.com/science/article/pii/S1361920921000687)

Broach, J., Dill, J., Gliebe, J.: Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transp. Res. A: Policy Pract. **46**, 1730–1740 (2012). https://doi.org/10.1016/j.tra.2012.07.005. (https://www.sciencedirect.com/science/article/pii/S0965856412001164)

Cheng, B., Misra, I., Schwing, A.G., Kirillov, A., Girdhar, R.: Masked-attention mask transformer for universal image segmentation, In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 1290–1299 (2022a)

Cheng, B., Misra, I., Schwing, A.G., Kirillov, A., Girdhar, R.: Masked-attention Mask Transformer for Universal Image Segmentation. (2022b). arXiv:2112.01527 [cs]

Cornet, Y., Lugano, G., Georgouli, C., Milakis, D.: Worthwhile travel time: a conceptual framework of the perceived value of enjoyment, productivity and fitness while travelling. Transp. Rev. **42**, 580–603 (2022). https://doi.org/10.1080/01441647.2021.1983067

Costa, M., Marques, M., Azevedo, C.L., Siebert, F.W., Moura, F.: Which cycling environment appears safer? Learning cycling safety perceptions from pairwise image comparisons. Trans. Intell. Transport. Sys. **26**, 1689–1700 (2025). https://doi.org/10.1109/TITS.2024.3507639

Costa, M., Marques, M., Siebert, F.W., Azevedo, C.L., Moura, F.: Bridging the gap: a scoping review exploring the relationship between objective and subjective cycling safety. Transp. Rev. **45**, 993–1015 (2025). https://doi.org/10.1080/01441647.2025.2537200

de Freitas, L.M., Axhausen, K.W.: What influences cycling infrastructure preferences? A stated-preference survey. Transportation (2025). https://doi.org/10.1007/s11116-025-10624-7

Dubey, A., Naik, N., Parikh, D., Raskar, R., Hidalgo, C.A.: Deep Learning the City: Quantifying Urban Perception at a Global Scale. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) Computer Vision - ECCV 2016, pp. 196–212. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_12

Fan, Z., Feng, C.C., Biljecki, F.: Coverage and bias of street view imagery in mapping the urban environment. Comput. Environ. Urban Syst. **117**, 102253 (2025). https://doi.org/10.1016/j.compenvurbsys.2025.102253. (https://www.sciencedirect.com/science/article/pii/S0198971525000067)

Fuest, S., Batista, M., Berghöfer, F.L., Flesser, M., Grandhi, B.S., Spuhler, F., Sester, M., Vollrath, M.: I bet you feel safe! Assessing cyclists' subjective safety by objective scores. J. Urban Mobil. **4**, 100066 (2023). https://doi.org/10.1016/j.urbmob.2023.100066

Gao, M., Fang, C.: Decoding the impact of audiovisual street environment features on cycling volumes: insights from street view imagery and machine learning. Transp. Res. A: Policy Pract. **199**, 104586 (2025). https://doi.org/10.1016/j.tra.2025.104586. (https://www.sciencedirect.com/science/article/pii/S0965856425002149)

Griswold, J.B., Yu, M., Filingeri, V., Grembek, O., Walker, J.L.: A behavioral modeling approach to bicycle level of service. Transp. Res. A: Policy Pract. **116**, 166–177 (2018). https://doi.org/10.1016/j.tra.2018.06.006. (https://www.sciencedirect.com/science/article/pii/S0965856417304688)

Grüner, S., Vollrath, M.: Reaching Your Destination on Time - Route Choice Decisions of Different Commuter Cyclist Types. In: Stanton, N. (ed.) Advances in Human Aspects of Transportation, pp. 162–169. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-80012-3_20

Hankey, S., Zhang, W., Le, H.T.K., Hystad, P., James, P.: Predicting bicycling and walking traffic using street view imagery and destination data. Transp. Res. Part D: Transp. Environ. **90**, 102651 (2021). https://doi.org/10.1016/j.trd.2020.102651. (https://www.sciencedirect.com/science/article/pii/S1361920920308361)

Hardinghaus, M., Papantoniou, P.: Evaluating cyclists' route preferences with respect to infrastructure. Sustainability **12**, 3375 (2020). https://doi.org/10.3390/su12083375

Hess, S., Rose, J.M.: Can scale and coefficient heterogeneity be separated in random coefficients models? Transportation **39**, 1225–1239 (2012)

Hevia-Koch, P., Ladenburg, J.: Where should wind energy be located? A review of preferences and visualisation approaches for wind turbine locations. Energy Res. Soc. Sci. **53**, 23–33 (2019). https://doi.org/10.1016/j.erss.2019.02.010. (https://www.sciencedirect.com/science/article/pii/S2214629618306662)

Hou, Q., Huo, X., Leng, J., Cheng, Y.: Examination of driver injury severity in freeway single-vehicle crashes using a mixed logit model with heterogeneity-in-means. Physica A **531**, 121760 (2019). https://doi.org/10.1016/j.physa.2019.121760. (https://www.sciencedirect.com/science/article/pii/S0378437119310088)

Huber, S., Lindemann, P., Schröter, B.: Safety and bicycle route choice: To what extent do accident risk and perceived safety influence bicycle route choice? Transp. Eng. **18**, 100240 (2024). https://doi.org/10.1016/j.treng.2024.100240. (https://www.sciencedirect.com/science/article/pii/S2666691X24000150)

Ito, K., Biljecki, F.: Assessing bikeability with street view imagery and computer vision. Transp. Res. C: Emerg. Technol. **132**, 103371 (2021). https://doi.org/10.1016/j.trc.2021.103371. (https://www.sciencedirect.com/science/article/pii/S0968090X21003739)

Ito, K., Bansal, P., Biljecki, F.: Examining the causal impacts of the built environment on cycling activities using time-series street view imagery. Transp. Res. A: Policy Pract. **190**, 104286 (2024). https://doi.org/10.1016/j.tra.2024.104286. (https://www.sciencedirect.com/science/article/pii/S0965856424003343)

Ito, K., Kang, Y., Zhang, Y., Zhang, F., Biljecki, F.: Understanding urban perception with visual data: a systematic review. Cities **152**, 105169 (2024). https://doi.org/10.1016/j.cities.2024.105169. (https://www.sciencedirect.com/science/article/pii/S0264275124003834)

Ito, K., Quintana, M., Han, X., Zimmermann, R., Biljecki, F.: Translating street view imagery to correct perspectives to enhance bikeability and walkability studies. Int. J. Geogr. Inf. Sci. **38**, 2514–2544 (2024). https://doi.org/10.1080/13658816.2024.2391969

Ito, K., Zhu, Y., Abdelrahman, M., Liang, X., Fan, Z., Hou, Y., Zhao, T., Ma, R., Fujiwara, K., Ouyang, J., Quintana, M., Biljecki, F.: ZenSVI: an open-source software for the integrated acquisition, processing and analysis of street view imagery towards scalable urban science. Comput. Environ. Urban Syst. **119**, 102283 (2025). https://doi.org/10.1016/j.compenvurbsys.2025.102283. (https://www.sciencedirect.com/science/article/pii/S0198971525000365)

Keppner, V., Sieber, C.C., Freiberger, E., Kob, R., Krumpoch, S., Siebentritt, H.M.: Characteristics of older cyclists with self-perceived needs for improvement in cycling competence: SiFAr trial. Eur. Geriatr. Med. **14**, 595–602 (2023). https://doi.org/10.1007/s41999-023-00765-2

Kuiper, M.D.: The impact of safety on cycling route choice: An analysis of the impact of objective and subjective safety aspects on cycling route choice behaviour in the Netherlands. Master's thesis, Utrecht University / TU Delft / University of Twente / Wageningen University (2021). (GIMA programme)

Lu, C., Piatkowski, D., Östh, J., Kourtit, K., Nijkamp, P.: Modeling the impact of street-level built environment on cyclists' route choice using street view images and GPS data. Environ. Plan. B: Urban Anal. City Sci. (2025). https://doi.org/10.1177/23998083251405122

Lukawska, M.: Quantitative modelling of cyclists' route choice behaviour on utilitarian trips based on GPS data: associated factors and behavioural implications. Transp. Rev. **44**, 1045–1076 (2024). https://doi.org/10.1080/01441647.2024.2355468

Madarshahian, M., Balaram, A., Ahmed, F., Huynh, N., Siddiqui, C.K.A., Ferguson, M.: Analysis of injury severity of work zone truck-involved crashes in South Carolina for interstates and non-interstates. Sustainability **15**, 7188 (2023). https://doi.org/10.3390/su15097188

Meng, L.: Political economy and cycling infrastructure investment. Transp. Res. Interdiscip. Perspect. **14**, 100618 (2022). https://doi.org/10.1016/j.trip.2022.100618. (https://www.sciencedirect.com/science/article/pii/S259019822200080X)

Ogawa, Y., Oki, T., Zhao, C., Sekimoto, Y., Shimizu, C.: Evaluating the subjective perceptions of streetscapes using street-view images. Landsc. Urban Plan. **247**, 105073 (2024). https://doi.org/10.1016/j.landurbplan.2024.105073. (https://www.sciencedirect.com/science/article/pii/S0169204624000720)

Oja, P., Titze, S., Bauman, A., De Geus, B., Krenn, P., Reger-Nash, B., Kohlberger, T.: Health benefits of cycling: a systematic review. Scand. J. Med. Sci. Sports **21**, 496–509 (2011). https://doi.org/10.1111/j.1600-0838.2011.01299.x

Pucher, J., Buehler, R.: Making cycling irresistible: lessons from The Netherlands, Denmark and Germany. Transp. Rev. **28**, 495–528 (2008). https://doi.org/10.1080/01441640701806612

Quintana, M., Gu, Y., Liang, X., Hou, Y., Ito, K., Zhu, Y., Abdelrahman, M., & Biljecki, F. (2025). Global urban visual perception varies across demographics and personalities. *Nature Cities*, 1–15. https://doi.org/10.1038/s44284-025-00330-x

Ren, K., Jin, M., Liu, H., Gong, Y., Liu, Y.: Discovering Cyclists' Visual Preferences Through Shared Bike Trajectories and Street View Images Using Inverse Reinforcement Learning. (2024). arXiv:2409.03148 [cs]

Reynolds, C.C., Harris, M.A., Teschke, K., Cripton, P.A., Winters, M.: The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. Environ. Health **8**, 47 (2009). https://doi.org/10.1186/1476-069X-8-47. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2776010/)

Rossetti, T., Daziano, R.: How does self-assessed health status relate to preferences for cycling infrastructure? A latent class and latent variable approach. Transportation **50**, 913–928 (2023). https://doi.org/10.1007/s11116-022-10266-z

Rossetti, T., Guevara, C.A., Galilea, P., Hurtubia, R.: Modeling safety as a perceptual latent variable to assess cycling infrastructure. Transp. Res. A: Policy Pract. **111**, 252–265 (2018). https://doi.org/10.1016/j.tra.2018.03.019. (https://www.sciencedirect.com/science/article/pii/S0965856417309278)

Salesses, P., Schechtner, K., Hidalgo, C.A.: The collaborative image of the city: mapping the inequality of urban perception. PLoS ONE **8**, e68400 (2013). https://doi.org/10.1371/journal.pone.0068400

Selvaraju, R.R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., Batra, D.: Grad-cam: Why did you say that?, (2016). arXiv:1611.07450

Sener, I., Eluru, N., Bhat, C.: An Analysis of Bicyclists and Bicycling Characteristics: Who, Why, and How Much are they Bicycling? Transportation Research Record 2134, (2009)

Skov-Petersen, H., Bernhard, B., Thomas, L., Jacobsen, J.B.: How do cyclists make their way? - A GPS-based revealed preference study in Copenhagen. Int. J. Geogr. Inf. Sci. **32**, 1469–1484 (2018). https://doi.org/10.1080/13658816.2018.1436713

Song, Q., Huang, Y., Li, W., Chen, F., Qiu, W.: Unraveling the effects of micro-level street environment on dockless bikeshare in Ithaca. Transp. Res. Part D: Transp. Environ. **132**, 104256 (2024). https://doi.org/10.1016/j.trd.2024.104256. (https://www.sciencedirect.com/science/article/pii/S136192092400213X)

Sorensen, M., Mosslemi, M.: Subjective and Objective Safety: The Effect of Road Safety Measures on Subjective Safety among Vulnerable Road Users. Technical Report 1009/2009. Institute of Transport Economics (TOI). Oslo, Norway, (2009)

Stinson, M.A., Bhat, C.R.: Commuter bicyclist route choice: analysis using a stated preference survey. Transp. Res. Rec. **1828**, 107–115 (2003). https://doi.org/10.3141/1828-13

Terra, R.W., Garrido-Valenzuela, F., Cats, O., van Cranenburgh, S.: Understanding cycling route choice behaviour through street-level images and discrete choice models. (2025). https://doi.org/10.2139/ssrn.5473360. https://papers.ssrn.com/abstract=5473360

Tilahun, N.Y., Levinson, D.M., Krizek, K.J.: Trails, lanes, or traffic: valuing bicycle facilities with an adaptive stated preference survey. Transp. Res. A: Policy Pract. **41**, 287–301 (2007). https://doi.org/10.1016/j.tra.2006.09.007. (https://www.sciencedirect.com/science/article/pii/S096585640600108X)

Ton, D., Cats, O., Duives, D., Hoogendoorn, S.: How do people cycle in Amsterdam, Netherlands?: Estimating cyclists' route choice determinants with GPS Data from an Urban Area. Transp. Res. Rec. **2662**, 75–82 (2017). https://doi.org/10.3141/2662-09. (**publisher: SAGE Publications Inc**)

Uijtdewilligen, T., Ulak, M.B., Wijlhuizen, G.J., Geurs, K.T.: Effects of crowding on route preferences and perceived safety of urban cyclists in the Netherlands. Transp. Res. A: Policy Pract. **183**, 104030 (2024). https://doi.org/10.1016/j.tra.2024.104030

Useche, S.A., Montoro, L., Sanmartin, J., Alonso, F.: Healthy but risky: A descriptive study on cyclists' encouraging and discouraging factors for using bicycles, habits and safety outcomes. Transport. Res. F: Traffic Psychol. Behav. **62**, 587–598 (2019). https://doi.org/10.1016/j.trf.2019.02.014. (https://www.sciencedirect.com/science/article/pii/S1369847818306934)

van Cranenburgh, S., Garrido-Valenzuela, F.: Computer vision-enriched discrete choice models, with an application to residential location choice. Transp. Res. A: Policy Pract. **192**, 104300 (2025). https://doi.org/10.1016/j.tra.2024.104300. (https://www.sciencedirect.com/science/article/pii/S0965856424003483)

van Cranenburgh, S., Garrido-Valenzuela, F. (2026). A utility-based spatial analysis of residential street-level conditions a case study of Rotterdam, *Cities, 174*, 107066. https://doi.org/10.1016/j.cities.2026.107066

Vedel, S.E., Jacobsen, J.B., Skov-Petersen, H.: Bicyclists' preferences for route characteristics and crowding in Copenhagen - A choice experiment study of commuters. Transp. Res. A: Policy Pract. **100**, 53–64 (2017). https://doi.org/10.1016/j.tra.2017.04.006. (https://www.sciencedirect.com/science/article/pii/S0965856416305973)

Verhoeven, H., Van Hecke, L., Van Dyck, D., Baert, T., Van de Weghe, N., Clarys, P., Deforche, B., Van Cauwenberg, J.: Differences in physical environmental characteristics between adolescents' actual and shortest cycling routes: a study using a Google Street View-based audit. Int. J. Health Geogr. **17**, 16 (2018). https://doi.org/10.1186/s12942-018-0136-x

Wang, Z., Ito, K., Biljecki, F.: Assessing the equity and evolution of urban visual perceptual quality with time series street view imagery. Cities **145**, 104704 (2024). https://doi.org/10.1016/j.cities.2023.104704. (https://www.sciencedirect.com/science/article/pii/S0264275123005164)

van Wijnen, J.H., Verhoeff, A.P., Jans, H.W., van Bruggen, M.: The exposure of cyclists, car drivers and pedestrians to traffic-related air pollutants. Int. Arch. Occup. Environ. Health **67**, 187–193 (1995). https://doi.org/10.1007/BF00626351

Winters, M., Babul, S., Becker, H.J.E.H., Brubacher, J.R., Chipman, M., Cripton, P., Cusimano, M.D., Friedman, S.M., Harris, M.A., Hunte, G., Monro, M., Reynolds, C.C.O., Shen, H., Teschke, K.: Safe cycling: How do risk perceptions compare with observed risk? Can. J. Public Health **103**, S42–S47 (2012). https://doi.org/10.1007/BF03403834

Xiao, Z., Chen, X., Zhang, C., Gong, Y.: Effects of road greenery and sky openness on non-commuting cycling route choice: insights from online map navigation. Transportation (2025). https://doi.org/10.1007/s11116-025-10659-w

Ye, Y., Zhong, C., Suel, E.: Unpacking the perceived cycling safety of road environment using street view imagery and cycle accident data. Accid. Anal. Prev. **205**, 107677 (2024). https://doi.org/10.1016/j.aap.2024.107677

Yu, X., Ma, J., Tang, Y., Yang, T., Jiang, F.: Can we trust our eyes? Interpreting the misperception of road safety from street view images and deep learning. Accid. Anal. Prev. **197**, 107455 (2024). https://doi.org/10.1016/j.aap.2023.107455

Zhang, W., Buehler, R., Broaddus, A., Sweeney, T.: What type of infrastructures do e-scooter riders prefer? A route choice model. Transp. Res. Part D: Transp. Environ. **94**, 102761 (2021). https://doi.org/10.1016/j.trd.2021.102761. (https://www.sciencedirect.com/science/article/pii/S1361920921000651)

Zhang, K., Song, Q., Ma, H., Qiu, W., Li, M., Kim, I.: Synergistic role of audio-visual perceptions in promoting bikeshare for active travel. Transp. Res. Part D: Transp. Environ. **145**, 104806 (2025). https://doi.org/10.1016/j.trd.2025.104806. (https://www.sciencedirect.com/science/article/pii/S1361920925002160)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Koichi Ito is a PhD student in the Department of Architecture at the National University of Singapore and a visiting student in the CityAI Lab in the Faculty of Technology Policy and Management, Delft University of Technology. His research interests include street view imagery, urban analytics, geographic data science, and GeoAI.

Roos Terra is a researcher affiliated with the Transport and Logistics group in the Faculty of Technology, Policy and Management and the CityAI Lab at Delft University of Technology. Her research interests include stated choice experiments, computer vision-enriched discrete choice models, and the use of street-level imagery to understand cyclists' route choice behavior and preferences for the cycling environment.

Miguel Costa is a postdoctoral researcher in the Machine Learning for Smart Mobility group, Transport Division of the Department of Technology, Management and Economics at the Technical University of Denmark. His research interests include computer vision and machine learning for urban perception, objective and perceived cycling safety, intelligent transportation systems, and sustainable urban mobility.

Felix Wilhelm Siebert is an Associate Professor of Transport Psychology in the Department of Technology, Management and Economics at the Technical University of Denmark. His research interests include road safety, transport psychology, human-technology interaction, and the application of computer vision methods to analyze road user behavior, including cyclist and motorcyclist safety.

Francisco Garrido-Valenzuela is a postdoctoral researcher in the CityAI Lab, Transport and Logistics group, Faculty of Technology, Policy and Management, at Delft University of Technology. His research interests include computer vision, street-level imagery, representation learning, and urban perception, with a focus on integrating these data sources into behavioral and discrete choice models.

Carlos Lima Azevedo is an Associate Professor in the Transport Division of the Department of Technology, Management and Economics at the Technical University of Denmark. His research interests include the modeling and simulation of human mobility, smart mobility services, transport demand and behavior modeling, traffic simulation, and machine learning for transport.

Filip Biljecki is an Assistant Professor and Presidential Young Professor in the Department of Architecture and the Department of Real Estate at the National University of Singapore, where he leads the NUS Urban Analytics Lab. His research interests include urban analytics, geographic information science, 3D city models, street view imagery, GeoAI, and geographic data science.

Sander van Cranenburgh is an Associate Professor of Choice Modelling in the Transport and Logistics group, Faculty of Technology, Policy and Management, at Delft University of Technology, where he leads the CityAI Lab. His research interests include discrete choice modelling, machine learning for choice and travel behaviour, non-RUM models, and the use of unstructured data such as images and text to understand human decision-making.