TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Terra, R., Garrido-Valenzuela, F., Cats, O., & van Cranenburgh, S. (2026). Understanding cycling route choice behaviour through street-level images and discrete choice models. *Transportation Research Part A: Policy and Practice*, *214*, Article 105274. https://doi.org/10.1016/j.tra.2026.105274

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Transportation Research Part A 214 (2026) 105274



Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra





Understanding cycling route choice behaviour through street-level images and discrete choice models

Roosmarijn Terra [a,*], Francisco Garrido-Valenzuela [a], Oded Cats [b], Sander van Cranenburgh [a]

[a] Transport and Logistics group, Faculty of Technology Policy and Management, Delft University of Technology, The Netherlands
[b] Department of Transport & Planning, Faculty of Civil Engineering & Geosciences, Delft University of Technology, The Netherlands

ARTICLE INFO

Keywords:
Discrete choice models
Computer vision
Street-level imagery
Stated-choice experiment
Cycling route-choice
Cycling environment

ABSTRACT

This study examines cyclists' preferences regarding infrastructural and environmental attributes, including lane type, pavement, lane colour, and the presence of greenery, water, and buildings. To do so, we conduct an image-based stated route-choice experiment in which respondents evaluate alternatives that vary in cycling environment, travel time, and the number of traffic lights. For the survey design, we used over six thousand real-world street-level images of the cycling environment, providing a more realistic choice context for respondents than could be achieved with a traditional stated-choice experiment based on tabular data with text and numbers. We analyse these route-choice data using two complementary modelling approaches: traditional discrete choice models with interpretable environmental attributes derived from GIS data, and a recently proposed computer vision-enriched discrete choice model that integrates image-based information directly into the discrete choice modelling framework. Results show that cycling environments are important factors in explaining cyclist route choice behaviour, with cyclists preferring green areas and separated cycling lanes over urban areas and mixed-traffic roads. In addition, we use the developed models to map where the cycling network in Rotterdam, our case study, aligns —or conflicts— with cyclists' preferences. Finally, we demonstrate how the developed model can provide planners with ex-ante evidence on the prospective effectiveness of proposed infrastructure upgrades.

1. Introduction

Cycling has become a central element of sustainable transport strategies. It is widely promoted because of its benefits for health, accessibility, sustainability, and liveability (Pucher and Buehler, 2010). It is also recognised as a key measure for achieving the 2015 Paris Climate Agreement targets, including the goal of a 40% reduction in CO_2 emissions by 2030 (I&W, 2020; Union, 2023). A crucial way for governments to encourage cycling is by providing high-quality infrastructure and supportive environments. Well-designed cycling environments have been shown to shift short-distance trips from car to bicycle (Fosgerau et al., 2023), improve perceived safety (Guo et al., 2023), and reduce accident risk (Costa et al., 2024a). To design such environments effectively, policymakers need a clear understanding of cyclists' preferences, thus, how different attributes of cycling routes – such as surface quality, roadside greenery, and travel time – are valued, and what trade-offs cyclists are willing to make among them.

* Corresponding author.
E-mail addresses: roosmarijn.terra@hetnet.nl (R. Terra), F.GarridoV@tudelft.nl (F. Garrido-Valenzuela), O.Cats@tudelft.nl (O. Cats), S.vanCranenburgh@tudelft.nl (S. van Cranenburgh).

https://doi.org/10.1016/j.tra.2026.105274
Received 11 September 2025; Received in revised form 20 July 2026; Accepted 12 September 2026
Available online 29 September 2026
0965-8564/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Discrete Choice Models (DCMs) are the dominant framework for investigating cyclists' preferences. By estimating utility functions from choice data, they quantify the relative importance of competing attributes in choice behaviour (McFadden, 1974). Both Stated Choice (SC) data, where hypothetical cycling route scenarios are presented to respondents (e.g., (Bovy and Bradley, 1985; Stinson and Bhat, 2003; Sener et al., 2009; Majumdar and Mitra, 2018; Liu et al., 2020; Vedel et al., 2017; Rossetti et al., 2018, 2019; Hardinghaus and Papantoniou, 2020; Significance, 2022; Uijtdewilligen et al., 2024; Tarkkala et al., 2024; De Freitas and Axhausen, 2025)), and Revealed Preference (RP) data, which comprise observed real-world cycling routes choices (e.g., (Broach et al., 2012; Ton et al., 2017; Bernardi et al., 2018; Skov-Petersen et al., 2018)) have extensively been used. From these studies, it is evident that several characteristics significantly influence cyclists' behaviour. These include: (1) Cycling infrastructure, such as lane type, pavement type, road width, and road colour; (2) Built environment factors, like greenery, buildings, street lights, parking and land-use; (3) Traffic-related attributes, including car traffic intensity, speed limits, cycling intensity and the presence of buses; (4) Perceptual factors, such as perceived attractiveness, traffic safety, and social safety and (5) Route-specific attributes, including travel time, travel distance, number of turns, intersections, and traffic lights. Notably, many studies find that the presence of dedicated cycling infrastructure and travel time are among the most influential factors to cyclists' route choices (Tilahun et al., 2007; Sener et al., 2009; Vedel et al., 2017; Rossetti et al., 2018; Hardinghaus and Papantoniou, 2020; Tarkkala et al., 2024; Bovy and Bradley, 1985; Stinson and Bhat, 2003; Broach et al., 2012; Ton et al., 2017; Skov-Petersen et al., 2018).

The majority of SC studies rely on textual descriptions to convey attributes of the environment, such as "the width of the bicycle path", "amount of street greenery", or perceptual labels, such as "busy", and "safe" (e.g. (Bovy and Bradley, 1985; Stinson and Bhat, 2003; Sener et al., 2009; Majumdar and Mitra, 2018; Liu et al., 2020)). Several studies incorporate visuals, such as sketches, abstract illustrations, or manually edited images, typically using a small, fixed set of carefully designed illustrations presented uniformly across respondents (Vedel et al., 2017; Rossetti et al., 2018, 2019; Hardinghaus and Papantoniou, 2020; Significance, 2022; Uijtdewilligen et al., 2024; Tarkkala et al., 2024; De Freitas and Axhausen, 2025). While these studies aim to represent the cycling environment in line with the researcher's intentions, they are limited in several significant ways. First, simplified visuals and textual descriptions often cannot accurately represent the full complexity of real-world environments (Florio et al., 2023). While these representations aim to capture key elements such as bike infrastructure and traffic, they often omit important contextual features, such as surface texture, vegetation, and the overall sense of safety or aesthetics (Florio et al., 2023; Hevia-Koch and Klinge Jacobsen, 2019). Moreover, when such features are visually present, they are seldom incorporated as explicit parameters in the choice models (See Rossetti et al., 2018; Tarkkala et al., 2024, in which trees appear in the visualisations yet no corresponding "tree presence" coefficient is estimated). Simplified visuals and textual descriptions also place a cognitive burden on respondents, requiring them to mentally reconstruct the environment, which may result in misinterpretation and biased preference estimates (Childers et al., 1985; Pinker, 1988). Finally, the use of a limited set of images can make results sensitive to specific images shown, thereby undermining external validity and generalisability (Hevia-Koch and Klinge Jacobsen, 2019). Consequently, the understanding of cyclists' preferences for cycling environments remains incomplete, largely due to the way in which these environments are presented in stated-choice experiments.

Unlike text descriptions or abstract visualisations, street-level imagery captures much of the complexity of cycling environments and allows researchers to present more valid contexts. Moreover, the growing availability of such imagery, together with advances in computer vision, enables analysis at scale (Kirillov et al., 2023). Recognising this potential, a parallel strand of research has emerged that examines cyclists' *perceptions* of the cycling environment using street-level images. Typically, their approach is as follows: respondents are shown paired street-view images and asked which environment feels safer or more pleasant for cycling; the responses are then used to train machine learning models that assign perceptual scores to new unseen images (Dubey et al., 2016). The resulting cycling environment utilities are then mapped to reveal spatial patterns of cycling perception (Costa et al., 2024b; Ito and Biljecki, 2021). Moreover, using pixel-level segmentation (e.g., percentages of greenery, buildings, or traffic elements), they are able to identify which visual features drive these cycling perceptions (Zeng et al., 2024). A comprehensive overview of this line of research is provided by Zhang et al. (2024) and Biljecki and Ito (2021). However, although perceptions provide valuable insights, they cannot explain cyclists' route choice behaviour. Perceptions are subjective interpretations of the environment (Wade and Swanston, 2013). Unlike preferences (which are formally rooted in choice theory (Samuelson, 1938; Luce, 1959; Lancaster, 1966)), perceptions do not capture the trade-offs individuals are willing to make that ultimately drive cycling choices.

In sum, the cycling literature splits into two strands with complementary strengths and limitations. The discrete-choice strand quantifies cyclists' trade-offs but typically represents environments through text or stylised visuals, which limits validity. The perception strand exploits street-level imagery at scale but elicits subjective perceptions rather than preferences, and therefore does not identify the trade-offs that govern route-choice behaviour. As such, current literature provides limited guidance for reliably designing and validating cycling infrastructure that aligns with cyclists' preferences.

This study addresses this gap by eliciting cyclists' preferences through an SC experiment using real street-level images. In our experiment, respondents chose between route alternatives that varied in the visual environment (image), travel time, and the number of traffic lights. To construct the experiment, we assembled a large set of cyclist-perspective images from the Municipality of Rotterdam's cycling infrastructure database. We then estimate three discrete choice models: two conventional logit models with explicitly encoded environmental attributes and a computer vision-enriched discrete choice model that learns salient visual features end-to-end from the imagery (Van Cranenburgh and Garrido-Valenzuela, 2025). We apply the latter model to map where Rotterdam's cycling network aligns –or conflicts– with cyclists' preferences, and we simulate targeted design changes to quantify their effects on predicted utility and route-choice probabilities, providing planners with ex-ante evidence on the prospective effectiveness of proposed upgrades.

The remainder of this article is organised as follows. Section 2 describes our methods: data collection, the design of the stated-choice experiment, and the model specifications, including the conventional discrete-choice models and the computer-vision-enriched

variant. Section 3 presents the estimation results and the model application to our case study: the cycling network in Rotterdam, the Netherlands. Finally, Section 4 provides a conclusion and discusses avenues for further research.

2. Methods

This Section covers three sub sections. First, we detail the collection of the images and cycling environment attributes in Paragraph Section 2.1. Second, we describe the design of the stated-choice experiment in Paragraph Section 2.2. Third, we present the estimated models in Paragraph Section 2.3.

2.1. Data collection

In the following, we outline the process of data collection in two parts. First the collection and categorisation of street-level images for the SC experiment. Then, the construction of a tabular dataset of cycling environment attributes for modelling preferences.

2.1.1. Street-level image collection and analysis

We use street-level images in our SC experiment to represent the cycling environment. Accordingly, our data collection process begins with gathering street-level images taken from the cyclist's perspective. These are obtained from the municipality of Rotterdam's image database (Cyclomedia, 2023) and captured between late February and early April 2023, reflecting typical late winter conditions. The full image database contains approximately 600,000 instances, including images from motorways, intersections, and areas not accessible to cyclists. We restrict image sampling to network segments where cycling is suitable. Additionally, we collect information about the key attributes in each image to support meaningful trade-off analyses and achieve a diverse representation of cycling environments. To do so, we first retrieve the cycling network (Fietsersbond, 2020), which consists of segments defined as stretches between intersections. For each segment, we place a midpoint and retrieve the nearest image within a 4-metre radius, ensuring that only locations suitable for cycling are included. Segments shorter than 3 metres (typically intersections) are excluded, as are images located in pedestrian zones, on bridges, or in parking areas. This process yields approximately 13,000 panoramic 360-degree images. Subsequently, we convert the panoramas into 90-degree forward-facing images aligned with the cycling direction.

Next, we analyse and categorise the images according to three key attributes: cycling infrastructure, built environment, and speed limit. These attributes are selected based on findings from previous cycling route choice studies. Using the known geographic location of each image, we link multiple spatial datasets (available via Fietsersbond (2020), BGT (2024), Rijkswaterstaat (2024)) to the images, resulting in a single integrated dataset in which each image is associated with a cycling infrastructure type, built environment context and a speed limit. Table 1 presents the attributes and their corresponding levels. We categorise the images into the 15 most frequently occurring attribute combinations. Categories with very few images are merged with similar ones to avoid over-representation of rare cases. The full list of the 15 attribute combinations is provided in Appendix A. These 15 categories serve as the image attribute levels in the SC design, providing meaningful trade-offs as well as a balanced representation of cycling environments.

Table 1
Characteristics and corresponding levels.

Cycling Infrastructure Fietsersbond (2020)	Built Environment BGT (2024)	Speed Limit Rijkswaterstaat (2024)
Solitary cycling lane	Recreational area	Not applicable
Separated cycling lane	Residential area	30 km/h
Mixed-traffic road	Neighbourhood access	50 km/h
Cycling suggestion lane	Main road	
	Industrial area	

2.1.2. Cycling environment attributes

We enriched the image sample with structured attribute data to estimate marginal utilities for cycling environment attributes visible in the street-level images. We considered two approaches. First, a computer-vision segmentation pipeline (e.g., Mask2Former) could extract semantic classes, such as trees, buildings, cars, roads, etc. at a pixel level. However, such pipelines are constrained by a predefined label set and struggle to recover several quantities central to cyclists' route choices (e.g., pavement type, tram line). Second, we could fuse information from external Geographic Information System (GIS) sources that describe the built environment, cycling infrastructure and traffic conditions. Given the detailed public Geographic Information Systems available in the Netherlands, the latter approach offers greater attribute richness and measurement accuracy. Therefore, we take the GIS-fusion route.

Concretely, we spatially joined each geolocated image to multiple GIS layers to derive: (1) *cycling-infrastructure* features (lane type, pavement, lane colour); (2) *built-environment* features (trees, grass, water, industrial and other buildings); and (3) *traffic and environmental* features (posted car speeds, weather and brightness conditions). Timestamps associated with the images enabled alignment to historical weather records from Crossing (2024). Different spatial buffers were applied during the merge: large features (e.g., buildings, water) were included if present within 20 metres of the image location, while smaller features (e.g., grass, planting, parking, tram lines) used a 10 metre buffer. Trees were counted along the length of the corresponding cycling segment. This means that the reported tree counts should be interpreted as indicators of greenery rather than as exact visual counts observable in images. The

resulting database comprises approximately 13,000 cyclist-perspective images, each linked to the cycling-environment attributes used for subsequent model estimation. Table 2 lists the full attribute set and provides example values for three images; the corresponding images are shown in Fig. 1.

Table 2
Collected attributes, linking method, and values for three example images.

Attribute	Levels	Source	Linking method	Example images		
				IMG1	IMG2	IMG3
Built Environment						
Trees	Count	BGT	Count within 20 m buffer	0	2	4
Grass	Yes/No	BGT	Present within 10 m	No	Yes	Yes
Plants	Yes/No	BGT	Present within 10 m	No	Yes	Yes
Water	Yes/No	BGT	Present within 20 m	No	Yes	No
Buildings	Yes/No	BGT	Present within 20 m	Yes	Yes	No
Industrial buildings	Yes/No	BGT	Present within 20 m	No	No	No
Cycling Infrastructure						
Infrastructure type	Categories	Fietsersbond	Initial network layer	Mixed-traffic road	Suggestion lane	Separated lane
Pavement type	Asphalt/Clinkers	Fietsersbond	Initial network layer	Asphalt	Asphalt	Asphalt
Parking	Yes/No	BGT	Present within 10 m	Yes	No	No
Colour	Red/Not red	BGT	Present within 10 m	Grey	Red	Red
Tram line	Yes/No	BGT	Present within 10 m	Yes	Yes	No
Traffic and Environment						
Car speed	30/50 km h^{-1}	Rijkswaterstaat	Present within 10 m	50	50	N/A
Weather	Sunny/Cloudy	Visual Crossing	Image timestamp	Cloudy	Cloudy	Sunny
Brightness	[0, 255]	Calculated	Computed from image	115	108	153



(a) IMG1 (b) IMG2 (c) IMG3

Fig. 1. Example images.

2.2. Stated choice experiment

This section outlines the setup of the SC experiment, including the selection of attributes and their levels, the experimental design and the layout and implementation of the survey. It also includes a descriptive analysis of the respondents' sample.

2.2.1. Selected attributes and ranges

In our SC experiment, respondents are presented with two hypothetical cycling routes. We conceptualise the cycling route alternatives using three attributes: the cycling environment (presented using a street-level image), the travel time and the number of traffic lights. These attributes are selected due to their consistent importance in cycling route choice behaviour and their relevance for policy and design interventions. Travel time has repeatedly been identified as a dominant factor in route choice decisions (Bovy and Bradley, 1985; Stinson and Bhat, 2003; Significance, 2022; Broach et al., 2012; Ton et al., 2017; Skov-Petersen et al., 2018). The number of traffic lights reflects interruptions to routes and has been shown to negatively affect cyclists' preferences (Sener et al., 2009; Vedel et al., 2017; Uijtdewilligen et al., 2024; Tarkkala et al., 2024; Broach et al., 2012; Ton et al., 2017; Skov-Petersen et al., 2018). While car speed and traffic intensity have also been found important in previous studies, they are not included as separate attributes in the design. As these factors are already partially conveyed through the images, including them would potentially create data congruency issues, see Sifringer and Alahi (2023). We use 8, 11, and 14 minutes for the travel time attribute, representing realistic durations for urban cycling trips; for traffic lights, we use 1, 2, and 3. The image attribute comprises 15 levels corresponding to the cycling environment categories defined in Section 2.1.1.

2.2.2. Efficient design

We employ an efficient design to maximise the information obtained from the SC experiment, resulting in more reliable parameter estimates for a given number of observations (Rose and Bliemer, 2009). Prior estimates for the model parameters are based on findings from Significance (Significance, 2022) and are reported in Appendix A. The design is created using Ngene software (ChoiceMetrics, 2024), generating two blocks of 15 choice tasks. The design includes labelled attributes, namely travel time, number of traffic lights, and the image category, with the latter reflecting a combination of infrastructure type, speed limit, and built environment type. In the experiment, for a given combination of infrastructure type, speed limit, and built environment type, images were randomly drawn from the image database.

2.2.3. Survey design

The survey was built and deployed with PixelSurvey (Garrido-Valenzuela et al., 2026), a YAML-configured, open-source platform built on Python's Dash framework (Plotly, 2024). The survey comprises three parts: 1) the SC experiment; 2) Perception questions involving the rating of images on traffic safety [1], social safety, and aesthetics, and 3) Questions about socio-demographic characteristics and cycling behaviour. Fig. 2 shows an example of a choice task from the experiment. Respondents are given the following instructions:

- Imagine you are cycling from your work, train station, school, or daily activity to your home.
- There are two cycling routes you can take.
- The routes differ only in **travel time**, the number of **traffic lights**, and the **cycling environment** shown in the photo.
- In all other aspects, such as the weather, the cycling routes are the same.
- **Which route would you choose?**
- You can assume the following:
  - You are cycling alone.
  - You are not in a hurry.
  - The photo gives a good idea of what the **entire route** looks like.
  - The travel time is the **total time** for the route, **including waiting at traffic lights.**
  - The weather is partly cloudy with no rain.

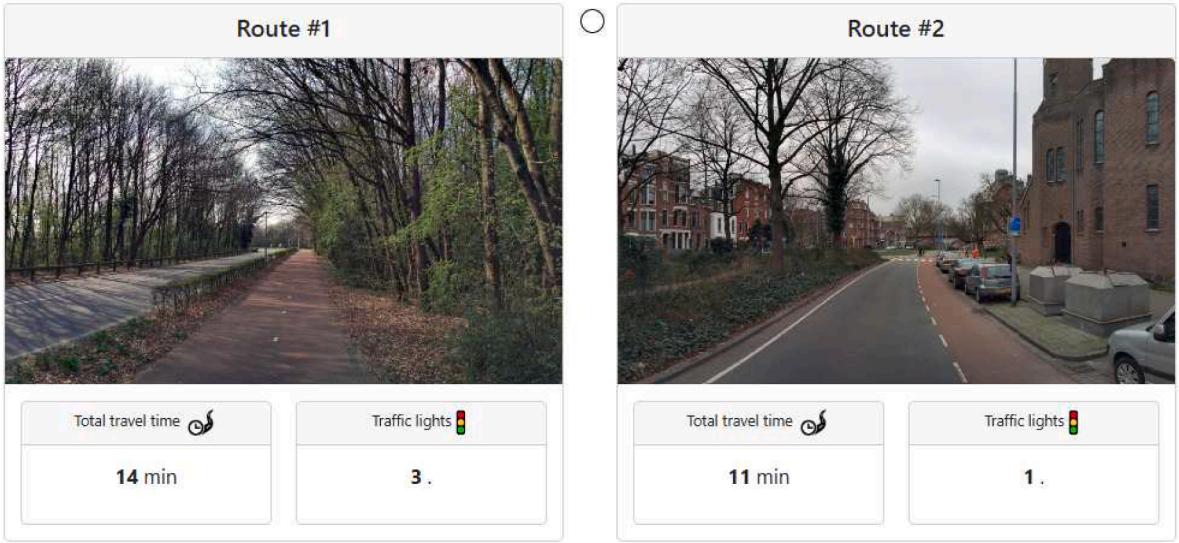


Fig. 2. Example of a choice task.

2.2.4. Survey implementation

The survey was conducted between late September and early October 2024. Respondents were recruited through a panel company and were financially compensated (Cint, 2024). The respondents were aged 18 years and above and regularly engaged in cycling. To ensure that the sample was broadly representative of the Dutch population aged 18 and above, respondent entry quotas were applied based on gender, age group, and region, using distribution targets taken from national statistics (CBS, 2024a). Quotas were enforced at screening: once a gender-age-region group reached its target, additional candidates in that group could not proceed.

[1] The same data have recently been used to investigate how cyclists trade perceived safety against travel time and the number of traffic lights (Ito et al., 2026).

Thirteen respondents who completed more than half of the choice tasks in under three seconds were excluded to ensure sufficient engagement with the choice tasks. In total, 746 valid responses were collected, corresponding to 11,190 observed choice tasks and involving 6484 unique images [2].

The dataset was split into a training set and a test set to reduce the risk of overfitting. Respondents were randomly assigned at the outset to either the training set (80%) or the test set (20%). Each set was linked to a disjoint pool of images, ensuring that no image shown to the training set appeared in the test set. This respondent-level, image-disjoint split prevents data leakage, i.e., models memorising specific images or respondents rather than learning generalisable visual attributes. The training set is used to estimate the parameters of the models, while the test set is held out during training and used for evaluation. Comparing model performance between the training and test sets provides an indication of the extent to which the model generalises to unseen data and helps to detect potential overfitting.

Each respondent was presented with 15 choice tasks. The choice task order and the position of alternatives (left/right) were randomised to prevent potential ordering effects. To avoid repetition, respondents did not see the same image more than once. Prior studies suggest that a high number of choice tasks may increase respondent cognitive burden and fatigue (Caussade et al., 2005; Rose et al., 2009). Although fifteen choice tasks may be considered relatively high, the use of images may have reduced cognitive burden. In fact, several respondents left comments and the end of the survey saying they thought the choice tasks were engaging. Nevertheless, potential fatigue effects cannot be fully ruled out.

2.2.5. Descriptive analysis

Respondents are asked about socio-demographic characteristics and their cycling habits. Overall, the sample is balanced in terms of age, gender, education, income and employment status and broadly reflects the national population. However, the sample over-represents individuals of Dutch origin (92%), which differs from the national population (72%) (CBS, 2024a). This may influence the generalisability of the results, particularly with regard to cycling behaviour and perceptions, which can vary across cultural backgrounds (Heinen et al., 2010). Approximately half of the participants identify cycling as their most frequent mode of transport. One-third of the respondents primarily cycle for commuting, while two-thirds cite other purposes such as recreation, shopping, or exercise. Previous research suggests that commuting cyclists tend to prioritise speed and directness, whereas recreational cyclists place more weight on the cycling environment (Bernardi et al., 2018). Respondents are also asked to rate the importance of the three route attributes in their decision-making. On average, the cycling environment (as presented by the image) is rated as the most important attribute (7.6), followed by travel time (6.1) and the number of traffic lights (5.5). While self-reported importance may not fully align with actual choice behaviour, these ratings offer a first insight into perceived priorities and suggest that the environment plays a significant role in route preferences.

2.3. Discrete choice models

2.3.1. Model overview

To examine cyclists' preferences regarding the cycling environment, we estimate three discrete choice models, all grounded in the Random Utility Maximisation (RUM) framework. Models 1 and 3 are estimated in multinomial logit form (MNL), while Model 2 is an Error Component Mixed Logit (EC-MXL). The utility specification of Model 2 is similar to Model 1, except that an individual-specific normally distributed error component is added to account for potential panel correlation. Both Models 1 and 2 derive the cycling environment attributes from the GIS-fusion route described in Section 2.1.2; In contrast, Model 3 uses an image encoder to extract relevant information about the cycling environment directly from the images shown in the survey. All models include an Alternative Specific Constant (ASC). Although the experiment is unlabelled, an ASC is included to capture any systematic tendency to favour one side of the screen:i.e., a left-right presentation effect (Boxebeld, 2024). For normalisation, the ASC of the left-hand side alternative is fixed to zero. The three models are described below in more detail:

1. **Model 1 (MNL with GIS cycling-environment attributes).** A standard linear-additive RUM-MNL model that explains cyclists' route choices using travel time, the number of traffic lights, and the GIS-derived cycling-environment attributes listed in Section 2.1.2. Because the cycling environment is represented by explicitly coded numerical attributes, the model yields interpretable marginal utilities for specific environmental features, e.g., trees, pavement type, infrastructure type, and buildings, and thus quantifies their association with route choice. The indirect utility function is given in Eq. 1, where x_{imn}^{gis} denotes the interpretable cycling-environment attributes derived from GIS data and ε_{in} denotes an i.i.d. Extreme Value Type I distributed error term.

$$U_{in} = asc_i + \beta_{\text{tt}} \cdot \text{tt}_{in} + \beta_{\text{tl}} \cdot \text{tl}_{in} + \sum_m \beta_m \cdot x_{imn}^{\text{gis}} + \varepsilon_{in} \tag{1}$$

2. **Model 2 (EC-MXL with GIS cycling-environment attributes).** This model extends Model 1 by adding an error component to one of the alternatives. As in Cabello et al. (2025), the purpose of this specification is to account for correlation in the unobserved utility across repeated observations from the same individual. As in the other models, the alternative-specific constant captures any mean left–right presentation effect. In addition, in this model, the error component, v_n, allows this presentation effect to vary across respondents, thereby capturing intra-respondent correlation in unobserved utility associated with the repeated presentation

[2] The processed data and code used in this study are available at https://github.com/TUD-CityAI-Lab/image-based-cycling-route-experiment.

of the alternatives.

$$U_{in} = asc_i + \beta_{\text{tt}} \cdot \text{tt}_{in} + \beta_{\text{tl}} \cdot \text{tl}_{in} + \sum_m \beta_m \cdot x^{\text{gis}}_{imn} + I_i \nu_n + \varepsilon_{in}, \quad I_i = \begin{cases} 0 & \text{if } i = 1 \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

3. **Model 3 (CV-DCM).** A computer-vision-enriched discrete choice model (CV-DCM) that integrates an image encoder into the utility function to ingest street-level images directly, rather than relying on GIS-coded attributes. We adopt the CV-DCM framework of Van Cranenburgh and Garrido-Valenzuela (2025), which has previously been published in this journal and is used here in an applied setting. Many attributes relevant to cyclists, including spatial openness, infrastructure continuity, and overall visual quality, are difficult to capture using GIS data yet are directly observable in street-level images. As a result, the CV-DCM is expected to outperform Model 1 in terms of predictive performance (e.g. log-likelihood). By assigning utilities directly to street-level images, the model captures cyclists' relative preferences for cycling environments; this enables the simulation of targeted infrastructure changes and the assessment of their effects on predicted utility, which we conduct in Section 3.2. However, these potential advantages come at the cost of interpretability, as the image-derived coefficients relate to latent embedding dimensions rather than semantically interpretable attributes. Moreover, unlike Model 2, the CV-DCM does not account for the panel structure by modelling correlation in unobserved utility across repeated observations from the same respondent. The indirect utility function of this model is given in Eq. 3, where z^{img}_{ikn} represents the image embedding of the street-level image. Since this model is newer and requires training (as opposed to estimation), the next section revisits this model briefly.

$$U_{in} = asc_i + \beta_{\text{tt}} \cdot \text{tt}_{in} + \beta_{\text{tl}} \cdot \text{tl}_{in} + \sum_k \beta_k \cdot z^{\text{img}}_{ikn} + \varepsilon_{in} \tag{3}$$

2.3.2. Computer vision-enriched discrete choice model

The computer-vision discrete choice model augments the traditional discrete-choice framework with a computer-vision component. For full methodological details, we refer the reader to Van Cranenburgh and Garrido-Valenzuela (2025). In this model, each street-level image is processed by a Data-Efficient Image Transformer (DeiT), which generates a compact embedding representation for each image. Specifically, each image is mapped to a $1 \times 1{,}000$ vector that summarises its salient visual features (Touvron et al., 2021). Through transfer learning, the image encoder is fine-tuned on the cycling route choice dataset so that the resulting embeddings, $\mathbf{z}_{in}$, capture visual features that are predictive of cyclists' route choices. These embeddings enter the utility function linearly alongside conventional numerical attributes such as travel time and the number of traffic lights, see the indirect utility function in Eq. 3.

The parameters associated with the image component can be retrieved after estimation; however, because the elements of the learned image embeddings lack explicit semantic meaning, these parameters cannot be interpreted as marginal utilities of identifiable environmental features. In contrast, the coefficients of the explicitly specified numerical attributes retain their conventional behavioural interpretation. Importantly, the CV-DCM is fully consistent with the random utility maximisation (RUM) behaviour, as the utility of alternative i depends exclusively on the attributes and street-level image associated with that same alternative. For practical and computational reasons, the model is trained in MNL form. Consequently, it shares the same limitations as Model 1: it does not explicitly account for the panel structure of the data and does not capture unobserved taste heterogeneity across individuals. Fig. 3 summarises the overall architecture.

Furthermore, to train the CV-DCM, we minimise the cross-entropy loss, which is equivalent to maximising the log-likelihood under the RUM framework (Aouad and Désir, 2025). We add an L2 regularisation penalty γ to mitigate overfitting. This is only applied to the feature-extractor weights w that produce the image feature map. We do not regularise the preference parameters–neither the numeric attribute coefficients β_m nor the image feature coefficients β_k–to avoid introducing bias in the preference estimates. The full objective is given in Eq. 4, where β denotes the preference parameters and w represents the computer vision weights.

$$w, \beta = \arg\min_{w,\beta} \left(\underbrace{\frac{1}{N} \sum_{n=1}^{N} \sum_{j=1}^{J} y_{nj} \log(P_{nj} | X_{nj}, S_{jn}, \beta)}_{\text{Cross-entropy loss}} + \underbrace{\gamma \sum_{r=1}^{R} w_r^2}_{\text{L2 regularisation}} \right) \tag{4}$$

We implement and train the CV-DCM using PyTorch. PyTorch is designed to build and train deep neural networks and facilitates GPU support, accelerating training. We conduct a hyperparameter search to identify a well-performing model configuration. In doing so, we systematically vary the optimisation algorithm, learning rate, batch size, and L2 regularisation strength. Table 3 summarises the hyperparameter combinations explored. The final model is selected based on the lowest cross-entropy loss on the test set, with early stopping and a patience of 5 epochs. The best-performing configuration employs stochastic gradient descent (SGD) with a learning rate of 1×10^{-6}, a batch size of 20, and an L2 weight-decay factor of 0.1, with the selected values highlighted in bold in Table 3.

2.3.3. Model evaluation

We assess model performance using standard discrete-choice evaluation metrics, including log-likelihood and McFadden's ρ^2. In addition, we report cross-entropy, defined as the negative mean log-likelihood, because it facilitates comparison across models evaluated on datasets of different sizes.

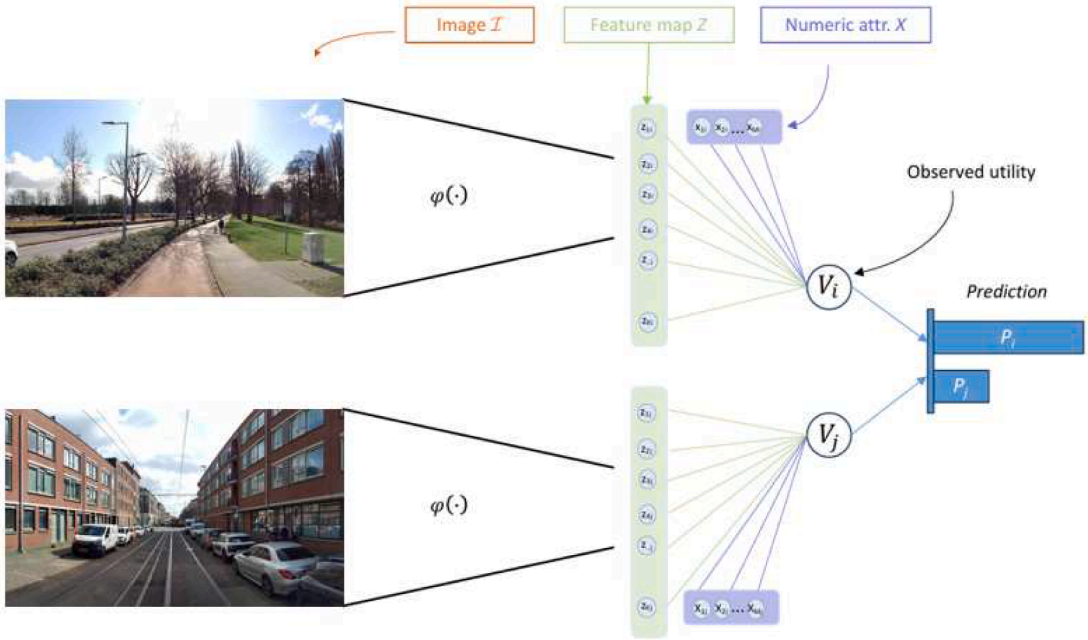


Fig. 3. Model structure of the CV-DCM.
Adapted from Van Cranenburgh and Garrido-Valenzuela (2025).

Table 3
Hyperparameter configurations explored for CV-DCM training.

Hyperparameter	Values Tested
Optimisation Algorithm	**SGD**, Adam
Learning Rate	$\mathbf{1 \times 10^{-6}}$, 1×10^{-5}, 1×10^{-4}
Batch Size	12, 16, **20**, 24
L2 Weight Decay (γ)	0, 0.05, **0.1**

Because Model 3 is a choice model that combines theory-driven and data-driven components (Van Cranenburgh et al., 2022), its performance cannot be evaluated reliably using in-sample measures alone. We therefore adopt a train-test split (see Section 2.2.4) and assess predictive performance on a hold-out sample. Hold-out evaluation provides a more credible indication of out-of-sample generalisation than inference based solely on estimation-sample statistics, such as likelihood ratio tests (Parady et al., 2021). This approach entails a trade-off, as a portion of the data must be reserved for evaluation rather than used for estimation, but it better reflects how the model would perform in practical applications.

The hold-out evaluation is used to compare the predictive performance of Model 3 against the benchmark models. Both Models 1 and 2 provide informative reference points. Model 1 offers the most direct comparison, as it shares the same MNL kernel as Model 3 and differs primarily in how the cycling environment is represented. Model 2 provides an additional benchmark that accounts for panel correlation through an error-component structure. We therefore report the predictive performance of all three models, allowing readers to assess the added value of Model 3 relative to either benchmark.

3. Results

This section comprises two parts. Section 3.1 reports the training and estimation results for the three models; Section 3.2 presents applications of our models to the city of Rotterdam, illustrating how our models' outputs can support planners with preference-based diagnostics of the cycling network.

3.1. Training and estimation results

Table 4 reports our estimation and training results for the three model specifications. We can make the following observations based on the performance indicators in Table 4. Model 1, the MNL model with GIS-coded cycling-environment attributes, attains the lowest model fit on the test set, with a LL of -1,323.4 and a corresponding ρ^2 of 0.091. Model 2, the EC-MXL, fits slightly better, increasing the model fit by 4 LL points. Model 3 improves model fit by 35 log-likelihood points relative to Model 2, achieving a

Table 4
Training and estimation results.

	Model 1			Model 2			Model 3		
Model type	MNL			EC-MXL			CV-DCM-MNL		
Number of parameters	16			17			86m		
Estimation/training time[a]	<1 sec			±40 min			±90 min		
Train set (*Resp* = 606 | *N* = 9,090)									
Log-Likelihood	−5,672.0			−5,665.6			−5,282.8		
ρ^2	0.100			0.101			0.162		
Cross-entropy	0.624			0.623			0.581		
Test set (*Resp* = 140 | *N* = 2,100)									
Log-Likelihood	−1,323.4			−1,319.1			−1,284.4		
ρ^2	0.091			0.094			0.118		
Cross-entropy	0.629			0.627			0.612		
	est	**s.e.**	**t-val**	**est**	**s.e.**	**t-val**	**est**	**s.e.**	**t-val**
Parameters									
asc_{right}	0.08	0.023	3.62	0.08	0.025	3.31	0.00	0.023	0.00
β_{tt}	−0.22	0.010	−22.67	−0.23	0.014	−16.30	−0.18	0.009	−21.07
β_{tl}	−0.31	0.018	−17.47	−0.32	0.022	−14.43	−0.27	0.017	−15.46
$\beta_{mixed\text{-}trafficroad}$	0.00	–fixed		0.00	–fixed				
$\beta_{separatedlane}$	0.40	0.054	7.56	0.41	0.055	7.46			
$\beta_{suggestionlane}$	−0.03	0.059	−0.54	−0.04	0.054	−0.65			
β_{bricks}	−0.11	0.043	−2.67	−0.12	0.042	−2.79			
$\beta_{parking}$	−0.07	0.038	−1.73	−0.07	0.039	−1.81			
β_{tram}	−0.23	0.111	−2.06	−0.24	0.107	−2.27			
β_{trees}	0.11	0.017	6.38	0.11	0.017	6.51			
β_{grass}	0.26	0.038	6.80	0.26	0.038	6.85			
β_{water}	0.18	0.038	4.60	0.17	0.039	4.45			
β_{plants}	0.02	0.037	0.44	0.02	0.038	0.57			
β_{house}	−0.26	0.049	−5.37	−0.27	0.053	−5.15			
$\beta_{industrial}$	−0.48	0.048	−10.09	−0.49	0.048	−10.30			
β_{sunny}	0.10	0.034	2.84	0.10	0.033	2.92			
$\beta_{brightness}$	0.00	0.001	5.83	0.00	0.001	5.52			
σ_{panel}				0.27	0.048	5.55			
Time-equivalent trade-off Traffic light [min]	1.38	0.103	13.30	1.38	0.128	10.79	1.49	0.099	15.12

[a] Models 1 and 2 were estimated on a laptop CPU, whereas Model 3 was trained on a mid-range GPU.

ρ^2 of 0.118 on the test set and is thus the best-performing model. This observation indicates that Model 3 can extract meaningful information from street-level imagery to explain choice behaviour. Moreover, its deep network architecture may capture nonlinearities and interactions among the visual attributes encoded in the images, enabling it to explain a larger share of the variance in cyclists' utilities derived from the cycling environment than Model 1 and 2 (in which all visual attributes enter the utility function in a linear form). Finally, Model 3 appears to exhibit moderate overfitting on the training set. However, some degree of overfitting is not necessarily problematic in this context, as the test set results show that the model still generalises well, and better than Models 1 and 2, to unseen choice tasks.

Next, we turn to the parameter estimates. The coefficients β_{tl} and β_{tt} represent the marginal utilities associated with the number of traffic lights and travel time, respectively. In line with behavioural intuition, both coefficients are negative in all models, indicating that cyclists prefer routes with fewer traffic lights and shorter travel times. Using these estimates, we can compute the mean time-equivalent trade-off: i.e., the additional cycling time a respondent is willing to accept in order to avoid one traffic light. This can be obtained as the ratio β_{tl}/β_{tt}. Based on Model 1 and Model 2, eliminating one traffic light is valued at 1.41 minutes of travel time. The CV-DCM yields a very similar estimate of 1.5 minutes, values that appear behaviourally plausible.

Turning to parameter estimates of the cycling environment attributes in Models 1 and 2, we see that the estimates are nearly identical between Models 1 and 2 and exhibit the expected signs (Table 4). The results show that cyclists strongly prefer separated cycling lanes over mixed-traffic roads. For example, the estimates from Model 1 imply that cyclists are willing to extend a trip of approximately 11 minutes by about 1.8 minutes in order to use a separated cycling lane rather than a mixed-traffic road, as implied by the ratio $\beta_{separatedlane}/\beta_{tt}$. Similarly, natural features such as trees, water, and grass are significantly preferred, whereas industrial areas, buildings and tramlines are less favoured by cyclists. Industrial areas generate the largest disutility ($\beta_{industrial}$ = -0.48), plausibly reflecting their low aesthetic quality and the prevalence of trucks and other motorised vehicles. Additionally, cyclists exhibit a clear preference for asphalt surfaces over brick paving. Although respondents were instructed to assume identical weather conditions across routes, the coefficient β_{sunny} is small but statistically significant, suggesting that respondents may still have been modestly affected by the weather condition in the image, despite the instruction to ignore it. Several candidate variables, namely cycling suggestion lane, plants, and parking, are statistically insignificant at the 5% level. Their lack of significance may come from indifference, but measurement issues could also play a role. For example, the plant variable covers only vegetation in public space, omitting greenery

in front gardens; cycling suggestion lanes typically occur on busy roads, which may not be fully captured by other GIS variables; and parking spaces do not necessarily imply the presence of parked vehicles.

Finally, we turn to the evidence for left–right presentation and panel effects. Models 1 and 2 exhibit small but statistically significant positive ASCs. Given that the experiment was unlabelled and alternatives were randomly assigned to the left or right side of the screen, this suggests that, on average, respondents choose the alternative presented on the right-hand side slightly more often than can be explained by the observed attributes included in these models. In addition, the statistically significant estimate of σ_{panel} in Model 2 provides evidence of a panel effect. It indicates that the unexplained tendency to favour one side of the screen is persistent across repeated choice tasks within respondents, while varying across respondents. Thus, some respondents systematically choose the right-hand alternative more often than can be explained by the observed attributes, whereas others systematically favour the left-hand alternative.

However, these findings should be interpreted with some caution. In Model 3, which provides the best fit, the estimated ASC is very small (1×10^{-5}) and statistically insignificant, and therefore does not support the notion that an average presentation effect is present. The discrepancy between the ASC estimates in Models 1 and 2, on the one hand, and Model 3, on the other, may have (at least) two complementary explanations. First, the ASC in Models 1 and 2 may compensate for residual variation that is not captured by the GIS-based specifications, whereas this variation is better explained by the image-based representation in Model 3. Second, the significant ASC may reflect sample-specific variation, meaning that the presentation effect is present in the training data but not in the test data by chance. This explanation is also consistent with the training procedure of Model 3. Because the training of Model 3 is guided by out-of-sample performance through early stopping, effects that are specific to the training data are unlikely to be retained if they do not improve generalisation.

To further assess the second explanation, we also estimated Models 1 and 2 directly on the test data (results not reported). In both models, the ASC flips sign ($asc_{right} = -0.02$) and is no longer statistically significant, which is consistent with the explanation that the estimated presentation effect is sample-specific. At the same time, when Models 1 and 2 are estimated on the full dataset, the ASC remains statistically significant. Taken together, these findings suggest that a modest average right-hand-side presentation effect may be present in the data, but that its magnitude is small and its contribution to out-of-sample predictive performance is limited. Similarly, while Model 2 provides statistical evidence of a panel effect through respondent-specific heterogeneity in the presentation effect, accounting for this panel structure yields only a modest improvement in fit and does not materially change the substantive conclusions.

3.2. Model applications to rotterdam

In this section, we apply our best-performing model, the CV-DCM, to illustrate how its outputs can support planners with preference-based diagnostics of the cycling network. Our case study is Rotterdam, the Netherlands' second-largest city, with approximately 660,000 residents (CBS, 2024b; Custers and Willems, 2024). The municipality aims to shift short trips from car to bicycle (Rotterdam, 2019), improve perceived safety (Rotterdam, 2023), and reduce traffic accidents (Fietsersbond, 2023). These objectives make Rotterdam a suitable setting for assessing how visual features of the cycling environment relate to cyclists' preferences.

We apply the trained CV-DCM to approximately 13,000 cyclist-perspective street-level images sampled across Rotterdam's cycling network (Section 2.1.1 details the sampling procedure). For each image, the model returns a systematic utility value, which we interpret as a preference-consistent indicator of the local cycling environment (higher values imply a greater predicted propensity to choose routes featuring such environments, *ceteris paribus*). We use these utilities to:

- Map the spatial distribution of the cycling environment utility, highlighting where the cycling network best and worst aligns with cyclists' preferences (Section 3.2.1);
- Identify the most salient visual attributes associated with the highest and lowest cycling environment utilities to highlight the recurring cues that distinguish preferred from non-preferred cycling environments.(Section 3.2.2);
- Examine associations between infrastructure types and built environment types and cycling environment utility (Section 3.2.3);
- Simulate design changes to provide planners with ex-ante evidence on the prospective effectiveness of proposed infrastructure upgrades. (Section 3.2.4).

3.2.1. Spatial distribution of the cycling environment utility

To reveal where Rotterdam's network best and worst aligns with cyclists' preferences, we map the cycling environment utility across all links of the cycling network. Because these utilities are relative, the colour scale highlights areas that outperform or underperform compared with the rest of the city. We then classify the utilities into five clusters with Jenks natural breaks, which minimise variance within clusters and maximise variance between clusters (Chen et al., 2013), see legend. Fig. 4 shows the resulting spatial distribution of the cycling environment utility for the cycling network of Rotterdam.

It is notable that the centre of Rotterdam is coloured red and orange (underperforming), while the surrounding regions are coloured yellow and light green (average to above average). The recreational areas and an area to the north-west, which is mostly composed of meadows, are represented by a dark green colour (overperforming). This contrast suggests that cyclists prefer green spaces over densely populated urban areas. A similar pattern is reported by Costa et al. (2024b) who found that cyclists perceive less urbanised areas with high levels of vegetation as safer.

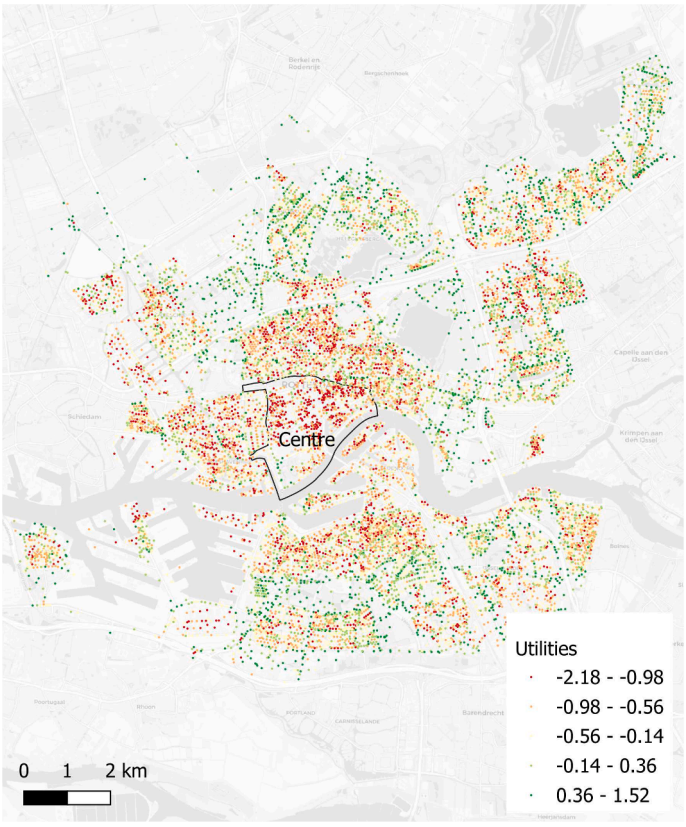


Fig. 4. Cycling environment utility for the cycling network of Rotterdam.

3.2.2. *Associations between cycling environment utility and visual attributes*

To understand what drives these cycling environment utilities, we explore the visual attributes that are associated with them. Specifically, we examine the visual attributes that associate with the highest and lowest cycling environment utilities. Accordingly, Fig. 5 shows nine random images from the highest and lowest utility clusters used to produce Fig. 4. We observe that the highest cycling environment utilities are associated with separated cycling lanes, green and open surroundings, and few buildings or traffic. In contrast, the lowest cycling environment utilities are associated with images of urban areas featuring buildings, vehicles, and narrow mixed-traffic roads. This pattern aligns with prior findings that cyclists prefer routes with greenery and separated lanes (Tarkkala et al., 2024; Uijtdewilligen et al., 2024; Rossetti et al., 2018).



(a) Images from highest cluster $(0.36 < V^{img} < 1.52)$. (b) Images from lowest cluster $(-2.18 < V^{img} < -0.98)$.

Fig. 5. Images sampled from the highest (left) and lowest (right) clusters based on cycling environment utilities.

3.2.3. Associations between cycling environment utility and infrastructure and built environment types

In this subsection, we examine how cycling environment utility varies across key dimensions of the cycling environment. To support behavioural interpretability, we analyse cycling infrastructure and speed limit jointly, and consider built-environment context separately.

We group the images based on the type of cycling lane and the maximum allowed car speed on the road. Fig. 6 presents the associations between cycling environment utility and cycling infrastructure (including speed limit) in a boxplot, ordered by increasing mean utility. Utilities are substantially higher for separated cycling lanes, whereas the difference between mixed-traffic roads and cycling suggestion lanes is marginal. This further supports the notion that the physical separation of the cycling lane from traffic is important to cyclists. The maximum car speed (30 km/h or 50 km/h) appears to be weakly associated with cycling environment utility. In contrast, previous studies have shown that higher car speed significantly reduces the attractiveness of a cycling route (Sener et al., 2009; Vedel et al., 2017; Rossetti et al., 2018; Hardinghaus and Papantoniou, 2020; Significance, 2022). A possible explanation is that respondents struggled to infer traffic speeds from the presented images.

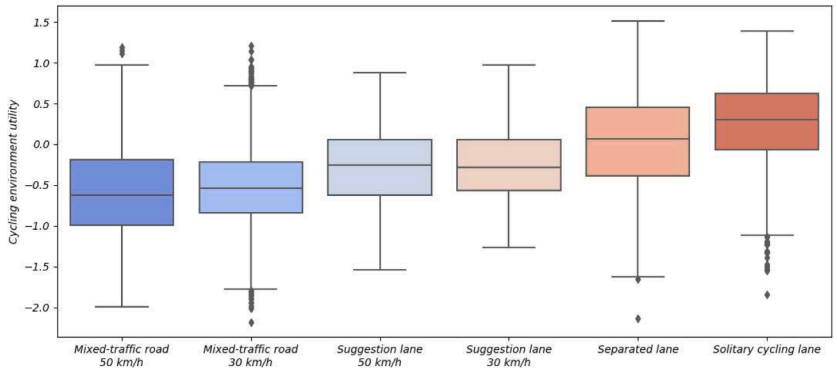


Fig. 6. Cycling environment utility by cycling infrastructure and car speed.

In addition, we examine how different built environment types are associated with cycling environment utility, see Fig. 7. The importance of built environment type to cycling route choice behaviour has rarely been studied; to the best of the authors' knowledge, only one study has examined this explicitly (Tarkkala et al., 2024). Indeed, most research relies on greenery indicators to represent the built environment. Fig. 7 shows that recreational areas, which characteristically combine separated cycling lanes with greenery, are associated with the highest utilities. These are followed by main roads and neighbourhood access roads. Residential and industrial areas are associated with the lowest utilities. Our results align with Tarkkala et al. (2024), who report higher disutility for local streets than for main roads, and with earlier evidence that cyclists strongly favour green surroundings (Vedel et al., 2017; Liu et al., 2020; Hardinghaus and Papantoniou, 2020).

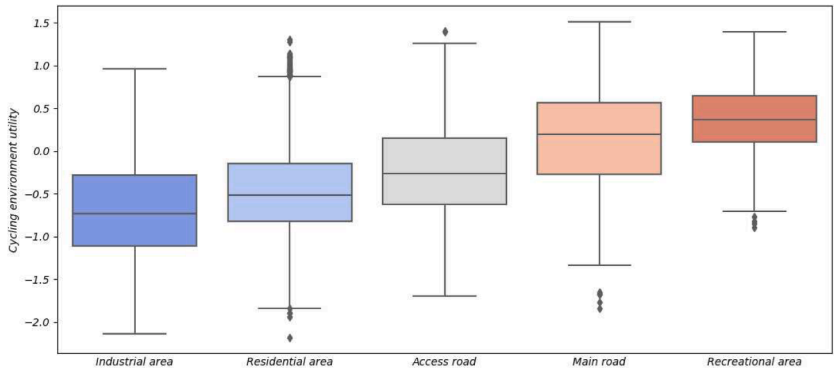


Fig. 7. Cycling environment utility by built environment type.

3.2.4. Ex-ante evaluation of cycling environment interventions

The municipality of Rotterdam has earmarked several roads for upgrades, including the road shown in the centre of Fig. 8. To examine how specific cycling environment design interventions affect cyclists' utility derived from the cycling environment in a route choice situation, we generated four synthetic variants of the baseline street-level image using FLUX.1 API (PiAPI, 2025). Fig. 8 displays five images (from left to right): (1) a cycling lane with a different surface colour, (2) the street without cars, (3) the baseline image, (4) the cycling lane removed, and (5) an alternative pavement type. All other visual elements are held constant. These images

are then evaluated with the trained CV-DCM. For each image, the figure reports its assigned utility (ranked from high to low) and, in the top-left corner, the estimated choice probability –defined as the probability of selecting the given cycling environment over the baseline for a ±11-minute cycling trip, all else being equal.

The colour of a cycling lane appears to matter. Changing it from grey to red increases the utility from -0.28 to −0.01, an improvement of 0.27. Based on the estimated travel-time coefficient ($\beta_{tt} = -0.18$), this translates into cyclists being willing to take a detour of one and a half minutes longer on an 11-minute trip in order to use a red lane instead of a grey one in the same setting. In direct comparison with the baseline, about 57% of cyclists would opt for the red-lane route. Similar colour swaps in other images consistently produce higher utilities, reinforcing the indication that lane colour matters to cyclists. This aligns with previous work: Dutch operational studies link red surfacing to better safety and smoother traffic flow (De Hoop, 2024), while stated-preference experiments show red lanes are judged safer and more appealing than other colours (Vera-Villarroel et al., 2016). Karlsen and Fyhri (2020) add a nuance, observing that colour preference can depend on users' familiarity with coloured cycling infrastructure. Taken together, both our model results and the literature suggest that lane colour can be an effective, low-cost design intervention for improving the perceived attractiveness of cycling environments to cyclists.

The absence of cars increases the cycling environment utility by 0.19 relative to the baseline version with cars. With the estimated travel-time coefficient ($\beta_{\text{tt}} = -0.18$), this difference translates into a willingness to accept roughly 1 minute of extra travel on an 11-minute trip in order to ride on a car-free road. The finding echoes the estimation results in Table 4 (Model 2), where the presence of parking, an indirect proxy for cars, decreases route utility. It is also consistent with the study of Uijtdewilligen et al. (2024), who find that traffic is a dominant factor in cycling route choices.

The utility difference between a marked lane and its absence is 0.10, indicating a modest preference for grey cycling suggestion lanes. This result also demonstrates that the CV-DCM can capture fine-scale surface markings, such as the narrow white stripe of a suggestion lane, with the feature extractor successfully identifying these elements in street-level imagery. Additionally, a substantial utility difference of 0.34 emerges between asphalt and brick pavement. This aligns with the findings of Kaparias et al. (2024), who report that cyclists favour asphalt over other pavement types. It is also consistent with the disutility estimated for brick surfaces by Model 2 as shown in Table 4. Taken together, the results indicate that cyclists prefer red lanes over grey, value car-free streets, and favour smooth asphalt surfaces, while subtle lane markings offer little added appeal. It also demonstrates the CV-DCM's value as a design-assessment tool, providing planners with quantitative feedback on cycling interventions prior to making any investments.



Fig. 8. Baseline image and four generated variants with their cycling environment utilities.

4. Conclusion

This study investigated cyclists' preferences for different cycling environments, focusing on features such as lane type, pavement type, and greenery. To this end, we conducted a novel image-based stated route choice experiment in which respondents evaluated alternatives that varied in cycling environment, travel time, and the number of traffic lights. Street-level imagery was employed to represent the cycling environment, with more than 6000 images used in the experiment.

We analysed the resulting choice data with both traditional discrete choice models and recently developed computer vision-enriched discrete choice models that integrate visual features extracted from images into the discrete choice modelling framework. Our findings demonstrate that the cycling environment is a significant determinant of cycling route choice behaviour. In particular, cyclists show a strong preference for separated cycling lanes and are willing to accept an additional 1.8 minutes of travel on an 11-minute trip to avoid mixed-traffic roads.

Finally, we illustrated how our models can provide planners with actionable insights. We mapped the extent to which Rotterdam's cycling network aligns–or conflicts–with cyclists' preferences and simulated targeted cycling environment design interventions to evaluate their ex-ante impact on cycling environment utility. These results highlight the potential of combining stated choice experiments with computer vision methods to support evidence-based planning for more attractive and cyclist-friendly cities.

Several limitations need to be acknowledged. While the use of real-world images in SC experiments offers significant advantages, it also raises important questions. For example, it is plausible that the visual representation of the cycling environment disproportionately captures attention, potentially diverting focus from numerically presented attributes (e.g., travel time and traffic lights in our study) and ultimately leading to biased estimates (Murwirapachena and Dikgang, 2022; Guilfoos et al., 2023; Li and Camerer, 2022). Future research could further investigate this hypothesis, e.g. through an A/B testing framework. Moreover, although images greatly

enhance respondents' understanding of the choice context compared to text and numerical descriptions alone, the presented scenarios may still deviate from real-world conditions. In our case, single images are used to represent entire routes, whereas in reality, the cycling environment can vary significantly along the route. Future studies could address this limitation by incorporating a series of images or videos, perhaps even with audio, to provide a more realistic representation of the choice situation.

Furthermore, the study is conducted in the Dutch context, where cycling is highly prevalent, infrastructure quality is generally high, and cycling is commonly perceived as a safe mode of transport. These context-specific conditions may influence the relative importance of cycling environment features identified in this study. In other settings, cycling decisions may be shaped more strongly by factors such as perceived insecurity or accident risk (e.g., Gutiérrez et al., 2021, 2020), potentially leading to different associations between cycling environment features and route choice. In addition, the computer vision model is trained exclusively on Dutch images. When applied to other countries, the model may therefore have difficulty recognising or correctly interpreting cycling environments that are uncommon in the Netherlands, limiting its transferability.

Finally, while the CV-DCM retains interpretability for numerical attributes, the computer vision component remains opaque. This limitation may be addressed by feature-learning approaches in which semantically meaningful image features are extracted and incorporated into traditional choice models; see, for example, Van Cranenburgh and Garrido-Valenzuela (2026). In addition, the CV-DCM in this paper is estimated with an MNL kernel, thereby ignoring panel effects that may be present in the data. At present, incorporating random effects through a mixed logit kernel is computationally challenging. Developing efficient methods to account for panel effects in CV-DCMs is, therefore, an important direction for future research.

CRediT authorship contribution statement

Roosmarijn Terra: Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization; **Francisco Garrido-Valenzuela:** Writing – review & editing, Supervision, Methodology, Data curation, Conceptualization; **Oded Cats:** Writing – review & editing, Supervision, Funding acquisition; **Sander van Cranenburgh:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Methodology, Funding acquisition, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During this work the authors used ChatGPT (OpenAI) to improve grammar and clarity in the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of this publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research, including the stated choice experiment and the subsequent work conducted while the first author was employed as a research assistant, was supported by funding from the Convergence Program – a collaboration between Delft University of Technology, Erasmus University Rotterdam, and the Municipality of Rotterdam. This work took place in the CityAI lab which is supported by the TU Delft AI initiative. No other specific grant from funding agencies was received for this research.

The authors thank the Municipality of Rotterdam for providing access to the street-level imagery used in this study. We also thank Kevin Su and Zoë Peters from the Municipality of Rotterdam for their involvement and contributions to the research. Finally we thank all participants of the stated choice experiment for their time and valuable input.

Data availability

The shareable part of the data is already made openly available.

Appendix A. Attribute combinations for the SC experiment

This appendix lists the 15 attribute combinations and their associated prior values used to categorise the street-level images and to construct an efficient design for the stated choice experiment (Table A.1).

Table A.1
Attribute combinations and priors for the 15 image categories. Priors adapted from Significance (2022).

Parameter	Prior value	Cycling infrastructure	Speed limit	Built environment
b_{time}	-0.2707			
$b_{trafficlights}$	-0.1330			
b_{cat0}	-1.002	Mixed-traffic road	50 km/h	Access road
b_{cat1}	-0.921	Mixed-traffic road	50 km/h	Industrial area
b_{cat2}	-0.850	Mixed-traffic road	50 km/h	Residential area
b_{cat3}	-0.720	Cycling suggestion lane	50 km/h	Access road
b_{cat4}	-0.663	Mixed-traffic road	30 km/h	Industrial area
b_{cat5}	-0.582	Mixed-traffic road	30 km/h	Main road
b_{cat6}	-0.512	Mixed-traffic road	30 km/h	Access road
b_{cat7}	-0.471	Cycling suggestion lane	30 km/h	Access road
b_{cat8}	-0.466	Mixed-traffic road	30 km/h	Residential area
b_{cat9}	-0.390	Separated cycling lane	N/A	Main road
b_{cat10}	-0.319	Separated cycling lane	N/A	Access road
b_{cat11}	-0.273	Separated cycling lane	N/A	Residential area
b_{cat12}	-0.116	Solitary cycling lane	N/A	Main road
b_{cat13}	-0.046	Solitary cycling lane	N/A	Residential area
b_{cat14}	0	Solitary cycling lane	N/A	Recreational area

References

Aouad, A., Désir, A., 2025. Representing random utility choice models with neural networks. Manag. Sci..

Bernardi, S., La Paix-Puello, L., Geurs, K., 2018. Modelling route choice of dutch cyclists using smartphone data. J. Transp. Land Use 11 (1). https://doi.org/10.5198/jtlu.2018.1143

BGT, 2024. Basisregistratie Grootschalige Topografie. https://www.kadaster.nl/zakelijk/registraties/basisregistraties/bgt.

Biljecki, F., Ito, K., 2021. Street view imagery in urban analytics and GIS: a review. Landsc. Urban Plan. 215, 104217. https://doi.org/10.1016/j.landurbplan.2021.104217

Bovy, P. H.L., Bradley, M.A., 1985. Route choice analyzed with stated-preference approaches. Transp. Res. Rec. (1037), 11–20. https://onlinepubs.trb.org/Onlinepubs/trr/1985/1037/1037-002.pdf.

Boxebeld, S., 2024. Ordering effects in discrete choice experiments: a systematic literature review across domains. J. Choice Model. 51, 100489. https://doi.org/10.1016/j.jocm.2024.100489

Broach, J., Dill, J., Gliebe, J., 2012. Where do cyclists ride? a route choice model developed with revealed preference GPS data. Transp. Res. A 46 (10), 1730–1740. https://doi.org/10.1016/j.tra.2012.07.005

Cabello, N., Galilea, P., de OrtúzarJ.D., 2025. Valuing improvements to bus universal accessibility for visually impaired users: a case study in santiago, chile. Res. Transp. Econ. 110, 101523. https://doi.org/10.1016/j.retrec.2025.101523

Caussade, S., de OrtúzarJ.D., Rizzi, L.I., Hensher, D.A., 2005. Assessing the influence of design dimensions on stated choice experiment estimates. Transp. Res. B 39 (7), 621–640. https://doi.org/10.1016/j.trb.2004.07.006

CBS, 2024a. Cbs. https://opendata.cbs.nl/statline/#/CBS/nl/.

CBS, 2024b. Statistieken Woonplaats Rotterdam. https://allecijfers.nl/woonplaats/rotterdam/.

Chen, J., Yang, S.T., Li, H.W., Zhang, B., Lv, J.R., 2013. Research on geographical environment unit division based on the method of natural breaks (jenks). Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XL-4/W3, 47–50. https://doi.org/10.5194/isprsarchives-XL-4-W3-47-2013

Childers, T.L., Houston, M.J., Heckler, S.E., 1985. Measurement of individual differences in visual versus verbal information processing. J. Consum. Res. 12 (2), 125–134. https://doi.org/10.1086/208501

ChoiceMetrics, 2024. Ngene. https://choice-metrics.com/download.HTML.

Cint, 2024. Cint. https://www.cint.com/.

Costa, M., Azevedo, C.L., Siebert, F.W., Marques, M., Moura, F., 2024a. Unraveling the relation between cycling accidents and built environment typologies: capturing spatial heterogeneity through a latent class discrete outcome model. Accid. Anal. Prev. 200, 107533.

Costa, M., Marques, M., Azevedo, C.L., Siebert, F.W., Moura, F., 2024b. Which cycling environment appears safer? learning cycling safety perceptions from pairwise image comparisons. arXiv preprint arXiv:2412.09835.

Crossing, V., 2024. Weather Data. https://www.visualcrossing.com/.

Custers, G., Willems, J.J., 2024. Rotterdam in the 21st century: from 'sick man' to 'capital of cool'. Cities 150, 105009.

Cyclomedia, 2023. Panoramafoto's Rotterdam Actueel. https://www.gis.rotterdam.nl/gisweb2/default.aspx.

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

De Hoop, B., 2024. Waarom Zijn Fietspaden in Nederland Rood? Het Verhaal van Tilrood. Accessed: 2024-12-09. https://www.dehoop-bouwgrondstoffen.nl/over-de-hoop/nieuws/waarom-zijn-fietspaden-in-nederland-rood-het-verhaal-van-tilrood.

Dubey, A., Naik, N., Parikh, D., Raskar, R., Hidalgo, C.A., 2016. 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. Springer International Publishing, Cham, pp. 196–212.

Fietsersbond, 2020. Data Fietsersbond. https://www.fietsersbond.nl/.

Fietsersbond, 2023. Wow! 1.600 Fietsongevallen per Jaar in Rotterdam: 'Terwijl we er Makkelijk Veel Kunnen Voorkomen'. https://indebuurt.nl/rotterdam/nieuws/verkeer/wow-1-600-fietsongevallen-per-jaar-in-rotterdam-terwijl-we-er-makkelijk-veel-kunnen-voorkomen~283785/.

Florio, P., Leduc, T., Sutter, Y., Brémond, R., 2023. Visual complexity of urban streetscapes: human vs computer vision. Mach. Vis. Appl. 35 (1), 7. https://doi.org/10.1007/s00138-023-01484-1

Fosgerau, M., Łukawska, M., Paulsen, M., Rasmussen, T.K., 2023. Bikeability and the induced demand for cycling. Proc. Natl. Acad. Sci. 120 (16). https://doi.org/10.1073/pnas.2220515120

Garrido-Valenzuela, F., Cats, O., van Cranenburgh, S., 2026. PixelSurvey: A YAML-based open-source framework for creating surveys. https://osf.io/preprints/psyarxiv/jkyhw_v1.

Guilfoos, T., Trandafir, S., Thomas, P., Uchida, E., Vogler, E., 2023. Visual representations in a choice experiment: valuing preferences for a local dam. Ecol. Soc. 28. https://doi.org/10.5751/ES-13898-280145

Guo, X., Tavakoli, A., Angulo, A., Robartes, E., Chen, T.D., Heydarian, A., 2023. Psycho-physiological measures on a bicycle simulator in immersive virtual environments: how protected/curbside bike lanes may improve perceived safety. Transp. Res. F 92, 317–336. https://doi.org/10.1016/j.trf.2022.11.015

Gutiérrez, M., Cantillo, V., Arellana, J., de OrtúzarJ.D., 2021. Estimating bicycle demand in an aggressive environment. Int. J. Sustain. Transp. 15 (4), 259–272. https://doi.org/10.1080/15568318.2020.1734886
Gutiérrez, M., Hurtubia, R., de OrtúzarJ.D., 2020. The role of habit and the built environment in the willingness to commute by bicycle. Travel Behav. Soc. 20, 62–73. https://doi.org/10.1016/j.tbs.2020.02.007
Hardinghaus, M., Papantoniou, P., 2020. Evaluating cyclists' route preferences with respect to infrastructure. Sustainability 12 (8). https://doi.org/10.3390/su12083375
Heinen, E., van Wee, B., Maat, K., 2010. Commuting by bicycle: an overview of the literature. Transp. Rev. 30 (1), 59–96. https://doi.org/10.1080/01441640903187001
Hevia-Koch, P., Klinge Jacobsen, H., 2019. Comparing offshore and onshore wind development considering acceptance costs. Energy Policy 125, 9–19. https://doi.org/10.1016/j.enpol.2018.10.019
Ito, K., Biljecki, F., 2021. Assessing bikeability with street view imagery and computer vision. Transp. Res. C 132, 103371. https://doi.org/10.1016/j.trc.2021.103371
Ito, K., Terra, R., Costa, M., Siebert, F.W., Garrido-Valenzuela, F., Azevedo, C.L., Biljecki, F., van Cranenburgh, S., 2026. Perceived traffic safety's impact on cycling route choice with street-level images. Transportation. https://doi.org/10.1007/s11116-026-10793-z
I&W, 2020. Klimaatbeleid. https://www.rijksoverheid.nl/onderwerpen/klimaatverandering/klimaatbeleid.
Kaparias, I., Miah, S., Chapman, A., Waterson, B., Milonidis, E., 2024. Using an instrumented bicycle to explore the responses of cyclists to pavement characteristics and defects. https://api.semanticscholar.org/CorpusID:274137543.
Karlsen, K., Fyhri, A., 2020. Is red the new black? A quasi-experimental study comparing perceptions of differently coloured cycle lanes. Front. Psychol. 11. https://doi.org/10.3389/fpsyg.2020.554488
Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment anything. arXiv preprint arXiv:2304.02643.
Lancaster, K.J., 1966. A new approach to consumer theory. J. Polit. Econ. 74 (2), 132–157. https://doi.org/10.1086/259131
Li, X., Camerer, C., 2022. Predictable effects of visual salience in experimental decisions and games. Q. J. Econ. 137, 1849–1900. https://doi.org/10.1093/qje/qjac025
Liu, Y., Yang, D., Timmermans, H. J.P., de Vries, B., 2020. Analysis of the impact of street-scale built environment design near metro stations on pedestrian and cyclist road segment choice: a stated choice experiment. J. Transp. Geogr. 82, 102570.
Luce, R.D., 1959. Individual Choice Behavior: A Theoretical Analysis. Wiley, New York.
Majumdar, B.B., Mitra, S., 2018. Analysis of bicycle route-related improvement strategies for two indian cities using a stated preference survey. Transp. Policy 63, 176–188. https://doi.org/10.1016/j.tranpol.2017.12.016
McFadden, D., 1974. The measurement of urban travel demand. J. Public Econ. 3 (4), 303–328. https://doi.org/10.1016/0047-2727(74)90003-6
Murwirapachena, G., Dikgang, J., 2022. The effects of presentation formats in choice experiments. Environ. Econ. Policy Stud. 24 (3), 421–445. https://doi.org/10.1007/s10018-021-00328-4
Parady, G., Ory, D., Walker, J., 2021. The overreliance on statistical goodness-of-fit and under-reliance on model validation in discrete choice models: a review of validation practices in the transportation academic literature. J. Choice Model. 38, 100257.
PiAPI, 2025. FLUX.1 API — State-of-the-Art Text-to-Image Generation. https://piapi.ai/flux-api. Accessed 27 June 2025.
Pinker, S. (Ed.), 1988. Visual Cognition. MIT Press, Cambridge, MA. https://archive.org/details/visualcognition00pink.
Plotly, 2024. Data Apps for Production. https://plotly.com/.
Pucher, J., Buehler, R., 2010. Walking and cycling for healthy cities. Built Environ. 36. https://doi.org/10.2148/benv.36.4.391
Rijkswaterstaat, 2024. Maxspeedcar. https://downloads.rijkswaterstaatdata.nl/nwb-wegen/.
Rose, J.M., Bliemer, M. C.J., 2009. Constructing efficient stated choice experimental designs. Transp. Rev. 29 (5), 587–617. https://doi.org/10.1080/01441640902827623
Rose, J.M., Hensher, D.A., Caussade, S., de OrtúzarJ.D., Jou, R.-C., 2009. Identifying differences in willingness to pay due to dimensionality in stated choice experiments: a cross country analysis. J. Transp. Geogr. 17 (1), 21–29. https://doi.org/10.1016/j.jtrangeo.2008.05.001
Rossetti, T., Guevara, A., Galilea, P., Hurtubia, R., 2018. Modeling safety as a perceptual latent variable to assess cycling infrastructure. Transp. Res. A 111, 252–265. https://doi.org/10.1016/j.tra.2018.03.019
Rossetti, T., Saud, V., Hurtubia, R., 2019. I want to ride it where i like: measuring design preferences in cycling infrastructure. Transportation 46 (3), 697–718. https://doi.org/10.1007/s11116-017-9830-y
Rotterdam, G., 2019. Fietskoers 2025. https://www.rotterdam.nl/fietsstad.
Rotterdam, G., 2023. Verkeersveiligheidsprikker. https://www.rotterdam.nl/verkeersveiligheidsprikker.
Samuelson, P.A., 1938. A note on the pure theory of consumer's behaviour. Economica 5 (17), 61–71. https://doi.org/10.2307/2548836
Sener, I., Eluru, N., Bhat, C., 2009. An analysis of bicycle route choice preferences in texas, US. Transportation 36, 511–539. https://doi.org/10.1007/s11116-009-9201-4
Sifringer, B., Alahi, A., 2023. Images in discrete choice modeling: Addressing data isomorphism in multi-modality inputs. arXiv preprint arXiv:2312.14724.
Significance, 2022. Values of Time, Reliability and Comfort in the Netherlands 2022. https://english.kimnet.nl/publications/publications/2024/04/03/new-values-of-travel-time-reliability-and-comfort-in-the-netherlands.
Skov-Petersen, H., Barkow, B., Lundhede, T., and, J. B.J., 2018. How do cyclists make their way? - a GPS-based revealed preference study in copenhagen. Int. J. Geogr. Inf. Sci. 32 (7), 1469–1484. https://doi.org/10.1080/13658816.2018.1436713
Stinson, M.A., Bhat, C.R., 2003. Commuter bicyclist route choice: analysis using a stated preference survey. Transp. Res. Rec. 1828 (1), 107–115. https://doi.org/10.3141/1828-13
Tarkkala, K., Vosough, S., West, J., Roncoli, C., 2024. Evaluating the influence of cyclists' route choices incorporation into travel demand modelling: a case study in greater helsinki. Transp. Res. Interdiscip. Perspect. 27, 101224. https://doi.org/10.1016/j.trip.2024.101224
Tilahun, N.Y., Levinson, D.M., Krizek, K.J., 2007. Trails, lanes, or traffic: valuing bicycle facilities with an adaptive stated preference survey. Transp. Res. A 41 (4), 287–301. https://doi.org/10.1016/j.tra.2006.09.007
Ton, D., Cats, O., Duives, D., Hoogendoorn, S., 2017. How do people cycle in amsterdam? estimating cyclists' route choice determinants using GPS data from an urban area. Transp. Res. Rec. 2662, 1–10. https://doi.org/10.3141/2662-09
Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., Jégou, H., 2021. Training data-efficient image transformers and distillation through attention. arXiv preprint arXiv:2012.12877.
Union, E., 2023. The importance of active mobility in the face of climate change. https://www.euroclima.org/en/recent-events-urban/articles-and-interviews/1981-world-bike-day-the-importance-of-active-mobility-in-the-face-of-climate-change.
Uijtdewilligen, T., Baran Ulak, M., Jan Wijlhuizen, G., Geurs, K.T., 2024. Effects of crowding on route preferences and perceived safety of urban cyclists in the netherlands. Transp. Res. A 183, 104030. https://doi.org/10.1016/j.tra.2024.104030
Van Cranenburgh, S., Garrido-Valenzuela, F., 2025. Computer vision-enriched discrete choice models, with an application to residential location choice. Transp. Res. A 192, 104300.
Van Cranenburgh, S., Garrido-Valenzuela, F., 2026. A utility-based spatial analysis of residential street-level conditions; a case study of rotterdam. Cities Forthcoming,
Van Cranenburgh, S., Wang, S., Vij, A., Pereira, F., Walker, J., 2022. Choice modelling in the age of machine learning-discussion paper. J. Choice Model. 42, 100340.
Vedel, S.E., Jacobsen, J.B., Skov-Petersen, H., 2017. Bicyclists' Preferences for route characteristics and crowding in copenhagen – a choice experiment study of commuters. Transp. Res. A 100, 53–64. https://doi.org/10.1016/j.tra.2017.04.006
Vera-VillarroelP., Contreras, D., Lillo, S., Beyle, C., Segovia, A., Rojo, N., Moreno, S., Oyarzo, F., 2016. Perception of safety and liking associated to the colour intervention of bike lanes: contribution from the behavioural sciences to urban design and wellbeing. PLoS ONE 11 (8), 1–15. https://doi.org/10.1371/journal.pone.0160399

Wade, N.J., Swanston, M., 2013. Visual Perception: An Introduction. Psychology Press, Hove, UK. 3rd edition. https://www.taylorfrancis.com/books/mono/10.4324/9780203082263/visual-perception-nicholas-wade-mike-swanston.
Zeng, Q., Gong, Z., Wu, S., Zhuang, C., Li, S., 2024. Measuring cyclists' subjective perceptions of the street riding environment using k-means SMOTE-RF model and street view imagery. Int. J. Appl. Earth Obs. Geoinf. 128, 103739. https://doi.org/10.1016/j.jag.2024.103739
Zhang, F., Salazar-Miranda, A., Duarte, F., Vale, L., Hack, G., Chen, M., Liu, Y., Batty, M., Ratti, C., 2024. Urban visual intelligence: studying cities with artificial intelligence and street-level imagery. Ann. Am. Assoc. Geogr. 114 (5), 876–897.