Cyclists in Motion
from data collection to behavioural models

Gavriilidou, A.

DOI
10.4233/uuid:b556b042-495f-4a75-a176-966c008074ee

Publication date
2021

Document Version
Final published version

Citation (APA)
Gavriilidou, A. (2021). *Cyclists in Motion: from data collection to behavioural models*. [Dissertation (TU Delft), Delft University of Technology]. TRAIL Research School. https://doi.org/10.4233/uuid:b556b042-495f-4a75-a176-966c008074ee

Cyclists in Motion

from data collection to behavioural models

Alexandra Gavriilidou

The research leading to this dissertation has received funding from the European Research Council under the European Union Horizon 2020 Framework Programme for Research and Innovation. It is established by the Scientific Council of the ERC Grant Agreement no. 669792 (Allegro).

Cyclists in Motion
from data collection to behavioural models

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus,
Prof.dr.ir. T.H.J.J. van der Hagen
chair of the Board for Doctorates
to be defended publicly on
Wednesday 3 February 2021 at 15:00 o'clock

by

Alexandra Gavriilidou
Master of Science in Civil Engineering
Delft University of Technology
born in Thessaloniki, Greece

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof.dr.ir. S.P. Hoogendoorn	Delft University of Technology, promotor
Dr.ir. W. Daamen	Delft University of Technology, promotor

Independent members:

Prof.dr. M. Hagenzieker	Delft University of Technology
Prof.dr. A. Seyfried	Institute for Advanced Simulation, Germany
Prof.dr. M. Bierlaire	École Polytechnique Fédérale de Lausanne, Switzerland
Prof.dr. R.L. Bertini	Oregon State University, USA
Dr. K. Watkins	Georgia Institute of Technology, USA
Prof.dr.ir. B. van Arem	Delft University of Technology, reserve member

TRAIL Thesis Series no. T2021/07, the Netherlands Research School TRAIL

TRAIL
P.O. BOX 5017
2600 GA Delft
The Netherlands
E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-282-7

Printed in the Netherlands

Αρχή ήμισυ παντός
(Well begun is half done)
-Plato

Cover illustration: Frank Auperlé
Cover design: Anna Paraschou
Flipbook design: Maria Stamatiadou & Periklis Panagiotou

Preface

When I was learning how to cycle, I remember wondering if there was a relation between the side from which the training wheel is removed first and the hand that the cyclist is most comfortable to keep off the handlebar. I had developed a theory and was collecting data to test my hypothesis, but my sample was so small I could not draw any conclusions. This dissertation does not provide an answer either, as it would have only served to please my curiosity and would have had a limited societal contribution. Instead, this dissertation is part of an ERC funded project, called Allegro, that aimed to unravel the travelling and traffic behaviour of pedestrians and cyclists in eight PhD projects. Mine focused on the traffic behaviour of cyclists. So, even though this dissertation answers different research questions, they do relate to the behaviour of cyclists and the research approach follows the principles I had adopted for that first research project, namely hypotheses formulation, data collection and analyses to test the hypotheses and draw conclusions. Despite the solid approach, this project would not have been a success without the support of my supervisors, the Allegro team, my friends and family. For this reason, I would like to express my gratitude to each and every one of them.

Serge, thank you for offering me this project, for composing this great Allegro team, for always keeping an eye on me and my well-being, for taking action when I needed it, as if you could read my mind, and of course for all the discussions we had about my research. Your enthusiasm, ideas and involvement in my project have been invaluable and kept me motivated. I enjoyed the time we spent looking together into the mathematics of the game theory models, and I really appreciate that you let me find my own way into the behavioural modelling and our great teamwork, also with Winnie, for the first ISTTT contribution.

Winnie, last year TRAIL awarded you the title of best supervisor of the year and the year before you were in the top three teachers of our MSc programme. To me, you have been the best teacher and the best supervisor every year and I am happy it could be officially acknowledged. You were genuinely interested in my research, always had a very thorough look at my documents and gave balanced and constructive feedback, challenging me to keep improving. I am also grateful to you for the opportunities you gave me regarding education tasks, by involving me in your course and guiding me in the supervision of master students. I learned a lot from you and I have greatly enjoyed the time we spent working together from brainstorming about new ideas and writing proposals to organising cycling experiments to supervising students and grading their reports. And that is only half of the story. The other half goes beyond your supervision role and I have admired you for keeping the two halves apart. Because during the last four years you have also been one of my best friends. You were the person I would come to when I needed to talk, and you would always listen and help me out. And I appreciate all the fun moments, from exploring new places, to attending football and tennis games, to cooking, playing games and watching movies. And a special thank you for being my corona buddy. I would never have made it through the lockdown without your company and our walks. So once again, thank you for your supervision and for your friendship.

Yufei, thank you for inspiring this project with your cycling experiments and for your constant support throughout my research. Your door was always open and you were happy to brainstorm with me whenever I needed that.

Haneen, thank you for your supervision and guidance during the early stages of my PhD project. You believed in me when I was losing hope, you encouraged me to pursue the research path I wanted and helped me define the boundaries of my research.

Dearest Allegri, it was my pleasure to be part of this team. It is said that doing a PhD is a lonely job, but with all of you around me, I never experienced it as such. We had a lot of fun during the outings and brainstorming sessions, working on the MOOC and travelling to conferences. A special word of thanks to my officemates. Marie-Jette, I was happy we could collaborate and organise this fun experiment together. Lara and Danique, thank you for all the talks behind closed doors. You showed me the way, gave me the confidence I needed and always turned my darkest hours several shades lighter. Florian, thank you for the hugs and for discussing your work with me. You offered the right amount of distraction and I really miss having you around in the office. Vincent thank you for all the laughs in the little time you spent in our office. I really miss your humour but I have enjoyed the closed windows in your absence. Tim, thank you for all the help with my coding, for thinking along and for letting me tease you every Thursday with the cookies! Giulia, thank you for the yoga exercises and healthy snacks.

I would also like to thank more colleagues, and friends, from the department for the lunches, the game nights and the outings, and over the last year for all the fun posts in the Stay-in-touch group. Thank you Panchamy, Tin and Boudewijn for the fun outings in The Hague. Thank you Pablo and Yihong for an unforgettable trip around California. Maria, thank you for all your advice, for sharing your energy and enthusiasm, for the walks in the corridors and the outings beyond that, but more importantly, thank you for your friendship. And of course, thank you and Chris for proof-reading my thesis. Alessandro and Elisa, thank you for all the dinners, the movies, for taking care of my plants when I was away and for the most epic afternoon working sessions under the sun in the Mekelpark. Priscilla, thank you for all your administrative support, but more importantly thank you for making me feel welcome into your office, hearing me out when I was struggling, and for sharing your office with me during the flex-working Fridays. Edwin and Peter, thank you for all the help preparing and setting up the experiment.

Outside of the department, I would like to thank all my friends for bearing with me, for distracting me, for hearing me out and advising me, for all the laughters and love that we have shared over the years. Luuk, Martijn, Marco, thank you for keeping in touch after the masters. Ozi, thank you for the roller-coaster ride and some awesome dinners. Anna, Faye, Violetta, Perikli, Elisa, Xristina, Niko, thank you for making every holiday break special, for the escape rooms, the board games, the dinners and movie nights. Anna, thank you for designing my cover and for being my best friend regardless of the distance separating us. Perikli, thank you for turning my flipbook idea into reality, for our unforgettable talks at Tony's and for all the laughters and fun memories on the Planet every summer. Gianni, Odussea, Alexandra, Xristina, thank you for keeping Fourka such a special place even after growing up.

Finally, a big thank you goes to my family, because without their unconditional love and support, none of this would have been possible. Ένα μεγάλο ευχαριστώ στις γιαγιάδες μου, για τα ανεπανάληπτα φαγητά τους και τις ώρες που περάσαμε μαζί παίζοντας χαρτιά, φτιάχνοντας πίτες και μιλώντας από κοντά ή από το τηλέφωνο. Λώρα, ευχαριστώ για όλες τις εξόδους στο στέκι μας και τις ατελείωτες συζητήσεις. Γιωργίτσα, είσαι η καλύτερη θεία του κόσμου, μέχρι και τον Ζιζούλη δέχτηκες να βάλεις στο σπίτι σου για χάρη μου. Και σου χρωστάω ακόμα ένα ταξίδι στο Άμστερνταμ. Πάνο, σε ευχαριστώ που είσαι μία σταθερά στη ζωή μου και που τους προσέχεις όλους όσο εγώ είμαι μακριά. Μαμά, μπαμπά, σας ευχαριστώ που στηρίζετε την κάθε μου τρέλα, που είστε πάντα δίπλα μου στο νου και στην ψυχή, και που με επισκέπτεστε τακτικά και εξερευνούμε μαζί την Ολλανδία. Λυπάμαι που δεν επιτρέπεται να παρευρεθείτε στην παρουσίαση του διδακτορικού μου, γιατί το μόνο που θέλω όταν τελειώσει η τελετή είναι να βρεθώ στην αγκαλία σας. Σας φιλώ από μακριά και ανυπομονώ να βρεθούμε ξανά, όταν όλα θα έχουν περάσει.

Alexandra
Delft, December 2020

Contents

List of Figures

List of Tables

Chapter 1

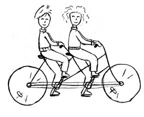
Introduction

This dissertation presents research on cycling behaviour, ranging from data collection and empirical findings to theories and models. The dissertation is based on research that also appears in four journal papers, which are brought together in a coherent story.

1.1 Background

Bicycles have been around for almost two centuries, but cycling has only began to be considered as a transport mode in the last few decades (Fishman & Cherry 2016; Pucher & Buehler 2012). This shift is attributed to the increased urbanisation which leads to traffic congestion and delays for car drivers, and also to the realisation that cars are not a sustainable solution, causing health and environmental problems. Several cities worldwide have, therefore, advocated the use of bicycles in the urban environment by providing incentives and systems such as shared bicycles (Baum 2008; Zhang et al. 2010; Avineri & Steven 2013; Dubuy et al. 2013). The benefits of such a shift in terms of public health and air quality are indisputable (Olde Kalter 2007; Heinen et al. 2010), yet its realisation needs to be supported by the urban design (Pucher & Buehler 2008).

So far, the design of cities has favoured the usage of motorised vehicles, which raises concerns regarding the safety and comfort of cyclists. In order to be able to accommodate large numbers of cyclists and ensure their safety, proper design of dedicated cycling infrastructure is needed. But what is proper? Answering this question requires understanding

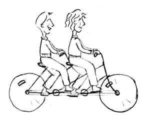


the behaviour of cyclists, how they interact with each other and how they make use of the infrastructure, as well as their needs and preferences. To this date, there is limited knowledge and understanding of these aspects of cycling behaviour.

Data and models are needed to shed light on these unknowns. Data provide empirical insights and are used to develop, calibrate and validate models. Traffic engineers then use these models to evaluate designs under varying traffic situations and make recommendations to policy makers and urban planners. The main hindrance to such developments and the required insights has been caused by the lack of empirical data. Even though a few models have been theoretically derived, they could not be calibrated or validated without data. This means that not only data are lacking, but also models describing cycling behaviour.

The type of data that is specifically needed to capture the movements and interactions of individual cyclists with each other and with the infrastructure is trajectories. In this context, a trajectory is a sequence in time of an individual's positions in space. Data regarding their needs and preferences can either be derived from their trajectories or collected by other means, such as surveys. For the calibration and validation of models that describe bicycle movements on dedicated cycling infrastructure, it is required to have trajectory data with high temporal resolution (less than a second) and spatial accuracy (a few centimetres). This level of detail makes it possible to observe how cyclists interact and how they use infrastructure. Examples of bicycle-to-bicycle interactions are cycling in groups and manoeuvres to overtake or to avoid collisions. The infrastructure utilisation encompasses, among others, the positioning of cyclists in a queue and the preference to take wide turns or cut corners.

In the last few years, several attempts have been made to collect cyclist trajectories. Researchers conducted bicycle experiments in which cyclists were asked to ride on circular tracks without overtaking (Andresen et al. 2013; Jiang et al. 2017). These experiments provided valuable insights into how cyclists follow each other and even led to the conclusion that the behaviour of cyclists is similar to that of car drivers and pedestrians when restricted to one lane (Zhang et al. 2014). However, the prohibition to overtake and the restriction of all movements into a single lane are not representative of bicycle motion, necessitating further data collection efforts.



The few models that have been derived, are based on behavioural models developed for cars. Even though their parameters were adjusted to reflect the lower speeds and smaller size of bicycles (Mallikarjuna & Rao 2009; Yao et al. 2009; Vasic & Ruskin 2012), it looks like these adjustments do not suffice. This is because they fail to capture the lane-free bicycle motion, the steering control and the physical effort required to balance. Apart from the behavioural shortcomings, these models focused on mixed traffic situations, i.e. bicycles and motorised traffic on the same road without physical separation (Oketch 2000; Mathew et al. 2012; Luo et al. 2015). In an urban environment that envisages to promote bicycle use and safety, infrastructure should be provided that is dedicated to cyclists (Wexler & El-Geneidy 2017). Thus, models are required that capture bicycle-to-bicycle interactions on dedicated cycling infrastructure.

To sum up, based on this background information two research gaps have been identified. First, there is a lack of cyclist trajectory data in different types of bicycle-to-bicycle interactions and traffic situations. Second, a model that describes cycling behaviour on dedicated cycling infrastructure has not yet been developed. Filling these gaps is the motivation of this dissertation.

1.2 Research questions

The aim of this dissertation is to develop a mathematical model that describes cycling behaviour, using cyclist trajectory data collected on dedicated cycling infrastructure. Cycling behaviour, in this context, covers the decisions and movements that cyclists make while cycling and interacting with other cyclists and with the infrastructure.

In order to meet this objective, the following research questions need to be answered:

1. Which modelling framework captures cycling decisions and movements?
2. Which are the key factors affecting the different decisions made while cycling?
3. Which datasets are needed to obtain the influence of the key factors on cycling decisions?

4. To what extent do the key factors influence the different decisions?

Given the collected datasets and developed models, empirical findings and behavioural insights are derived. These are used to make design recommendations for cycling infrastructure by answering the final research question of this dissertation:

5. Which design implications stem from the empirical findings and behavioural insights of the derived datasets and models?

1.3 Research approach

The research approach is set up such that the research questions can be answered and the objective be met.

The first research question poses the need for a modelling framework that captures cycling behaviour. In order to develop such a framework, the first step is to define which decisions are covered by the term cycling behaviour, and to understand how the decisions of the cyclists and their movements should be linked. To this end, existing definitions for the behaviour of car drivers and pedestrians were consulted and adjusted to match that of cyclists. The second step is to demarcate the decisions under consideration. This demarcation is based on an overview of traffic situations that cyclists encounter while cycling on dedicated infrastructure.

For each of these decisions, (different) key factors play a role. The second research question aims to identify these key factors. A literature study was performed to reveal the key factors for decisions that have been previously investigated. A survey was constructed to derive the rest, including some overlap with what is known from literature to check the extent to which they match and as such justify the predictive value of the survey.

The next step is to develop theoretical models that fit the framework and take into account the key factors. In this dissertation, discrete choice theory is used, because it is in line with the behavioural assumptions and it offers a consistent modelling approach for all decisions within the framework. More explanation relating to this choice can be found in chapter 2, where the framework and behavioural assumptions are introduced and discussed.

Before being able to apply these models in practice, they have to be calibrated and validated. As previously argued, bicycle trajectory data are needed for the calibration and validation, but are not yet available for all traffic situations (i.e. decisions). In order to acquire trajectory data on those situations that are missing, while ensuring the desired temporal accuracy and spatial resolution, a controlled cycling experiment is conducted. The design of the controlled experiment answers the third research question and its implementation generates the collected datasets. Further motivation for the performance of a controlled experiment as the data collection method is provided in chapter 3.

Using the collected datasets, discrete choice models are estimated. Their estimation process leads to calibrated models that are further validated by means of a simulation. The resulting models reveal the extent to which the different factors influence the cycling decisions. This answers the fourth research question and generates the final product of this dissertation, namely a cycling behaviour model.

Finally, recommendations are made for the design of cycling infrastructure. These are based on the empirical findings generated by processing the collected datasets and the behavioural insights gained by the calibrated models. With these design recommendations, the investigation of cyclist behaviour comes to a full circle and the dissertation fills the gaps identified in the background.

1.4 Contributions

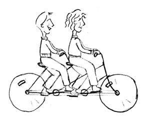
This dissertation achieves several contributions, both scientific and practical. In this section an overview is provided, while details can be found in the corresponding chapter(s).

The main scientific contribution is the development of a modelling framework for cycling behaviour at the operational level (chapter 2). A secondary contribution is the definition of the operational level for cycling behaviour. According to the definition, the operational behaviour level covers decisions and actions taken while cycling, which means that a destination has been selected, as well as a route through the network towards this destination. In the operational level, cyclists choose their path within the route, which relates to their positioning in the lane, and also the control dynamics of the bicycle, namely the pedalling and steering. The added value of this definition is that it marks the boundaries regarding the types of behaviour and decisions that the operational level encompasses. Based on this definition, the modelling framework follows a two-layer approach, which first divides the decisions into mental (path choices) and physical (control dynamics), and then represents the decision making process by linking the relevant decisions to each other.

In addition to the modelling framework, this dissertation produced a rich dataset of cyclist trajectories. A large-scale controlled cycling experiment was performed and captured high cyclist volumes and different types of interactions, such as overtaking, bidirectional traffic, merging and yielding. These data can be used by other researchers to study cycling behaviour and generate new models and insights. Apart from this scientific contribution, there is also a practical one in the performance of the experiment. The process to set it up is delineated in chapter 3, along with lessons learnt from implementing it, which can be used as a guide for future experiments studying operational behaviour.

Further contributions are to be found in the estimated discrete choice models. Firstly, the derived models capture different situations and operational decisions of cyclists. More specifically, the models describe the bicycle queue formation process at a signalised intersection (chapter 2) and the cyclist interactions at unsignalised intersections (chapter 5). The former shows how cyclists brake to a complete stop as well as their choice for a queuing spot. The latter describes yielding behaviour for cyclists who have to give priority and a regular cycling model for those cyclists that have priority at the intersection. These models contribute to the cycling behaviour modelling suite and can be used by researchers and practitioners to develop simulation models and investigate different scenarios, assess different situations and provide advice to policy makers. Secondly, the estimation of these models demonstrates the generalised application of the two-layer framework, leading the way for researchers to develop more models to capture different decisions within the framework.

The data analyses as well as the derived models generate valuable empirical findings and behavioural insights. Based on these, design recommendations are put forward (chapters 4 and 6) which constitute another practical contribution.



Last but not least, the societal relevance and contribution lies in the use of these products (data, models, design recommendations) by urban planners and policy makers to provide safe and comfortable cycling infrastructure.

1.5 Outline

This dissertation revolves around three pillars, namely models, data and design recommendations. These are decomposed into six elements and linked as shown in Figure 1.1. In order to derive mathematical models, a modelling framework is needed, as well as data. When data are not available, they need to be collected first and then analysed. The models together with the data analyses lead to behavioural insights, which are used to generate design recommendations. Each chapter in this dissertation combines several of these elements, as displayed in the figure.

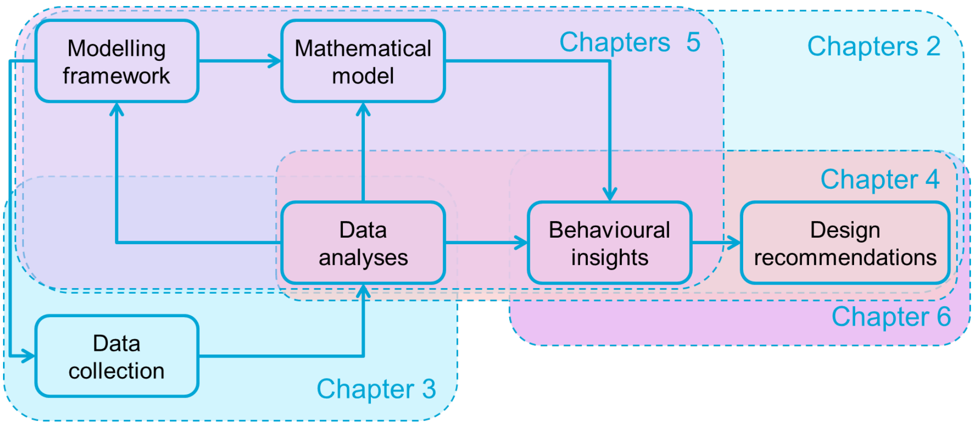


Figure 1.1: Elements covered per thesis chapter.

More specifically, the remainder of this dissertation contains five chapters. Chapter 2 presents the modelling framework that is proposed to capture cyclist decisions and movements. In this chapter, an application of the framework is also provided using an existing dataset (hence no data collection) of cyclists queuing at a signalised intersection. Discrete choice models are estimated, leading to behavioural insights for the queue formation upstream of red traffic lights. Based on these insights, recommendations are made to improve the queue formation. In order to be able to develop models for other situations, extensive data collection has been performed and is presented in chapter 3. This data collection includes the survey to derive the key factors, as well as the controlled experiment to obtain trajectories of different individuals in several traffic situations and interactions. Chapter 4 contains a first analysis performed on the experimental data and empirical findings regarding the traffic efficiency of bicycle T-junctions. Additionally, the use of lane marking for guidance of cyclists is evaluated and design recommendations are made. In chapter 5, the modelling framework is revisited and updated, and using the dataset from the controlled experiment, discrete choice models are estimated. These describe cycling decisions made upstream a bicycle unsignalised intersection and enhance the behavioural insights into operational cycling behaviour. The thesis closes with chapter 6, where conclusions are drawn from the behavioural insights and recommendations are made regarding the infrastructure design, the application of the developed models and future research in this field.

Chapter 2

Two-layer modelling framework for operational cycling behaviour

As explained in the introduction, operational cycling behaviour is greatly understudied. Moreover, a definition is lacking of what the operational behavioural level actually entails in terms of decision making. In this chapter, the operational level for cycling behaviour is defined and a two-layer modelling framework is proposed. The two layers stem directly from the definition, according to which the operational level consists of two intertwined processes, a mental and a physical process, each captured by the corresponding layer in the framework. The mental process refers to path choices made within a route, while the physical process refers to the bicycle control dynamics through pedalling and steering. The application of the framework is demonstrated by developing mathematical models that represent the cycling behaviour during one of the processes within the framework, namely the queue formation process of cyclists upstream of a red traffic light. This process includes selecting a queuing position (mental layer) and cycling towards it (physical layer). The proposed framework, the developed models and the behavioural findings on the bicycle queue formation process obtained by the application of this framework are the main contributions of this chapter. The framework is revisited in chapter 5 with another application at a bicycle crossing.

This chapter is based on the published article: Gavriilidou, A., W. Daamen, Y. Yuan, S. Hoogendoorn (2019a) Modelling cyclist queue formation using a two-layer framework for operational cycling behaviour, *Transportation Research Part C: Emerging Technologies*, 105, pp. 468–484.

2.1 Introduction

Though the interest in cycling in cities increases, research on bicycle traffic behaviour is still in its infancy. Insights into this behaviour, and understanding how cyclists interact with each other and make use of cycling infrastructure are crucial if cities are to be designed to accommodate large amounts of cyclists and ensure their safety. Since models can be used to evaluate different designs under varying traffic situations, this need for insights is linked to the need to create reliable and accurate models that can, for example, assess the capacity of intersections or predict the number of encounters on bi-directional cycle paths as a surrogate safety measure.

Research on how cyclists make use of the infrastructure is, however, limited. Among the few examples is Jiang et al. (2013), who studied the gap acceptance of cyclists against right-turning vehicular traffic at signalised intersections. They found that cyclists started decelerating when they are within 30 m from the stop line and that their acceptance of a gap depends on the speed of the cyclist and of the motorised vehicle, as well as the size of the available gap. Kucharski et al. (2019) observed the formation of multiple channels in queues at signalised intersections and found that the number of channels formed correlated with the length of the queue. Since they only looked at a single intersection, it is possible that the effect of other factors, such as the width of the cycle path, has not been identified. In line with this remark, the authors stressed the need for a bigger sample before a model could be formulated to describe the queue formation process.

More research effort has been put on modelling the bicycle control dynamics while riding and interacting with other road users, as several microscopic behavioural models have been developed. Early microscopic cyclist models made use of modelling paradigms developed for cars, such as Cellular Automata models, while adjusting their parameters to reflect the lower speeds of bicycles and their smaller size (Mallikarjuna & Rao 2009; Yao et al. 2009; Vasic & Ruskin 2012). However, the rules governing the movement between cells have not been adjusted to represent cycling behaviour. Another example is the car-following model that was derived for bicycle traffic by Andresen et al. (2013). Even though it was calibrated using empirical cyclist data, the model described single file bicycle flow which is generally not representative of cyclist movements on cycling infrastructure. In addition to these, models stemming from research on pedestrian dynamics were developed to model the microscopic cycling behaviour, such as social force models that determine the movement of cyclists based on attractive forces towards the desired destination and repulsive forces from obstacles and other traffic users, including other cyclists (Li et al. 2011; Liang et al. 2012; Huang et al. 2017). Utility-based models constitute another approach to describe pedestrian dynamics, but their application to cycling is so far scarce. Even though NOMAD (a game theoretical model developed to capture pedestrian dynamics (Hoogendoorn & Bovy 2002)) was adjusted for bicycle traffic by Twaddle (2017), these adjustments come with strong behavioural assumptions that make the model mathematically tractable and move away from the game theory, turning it into a social force model. At the same time, a game theoretical approach has been applied and was deemed plausible (Gavriilidou et al. 2017), but it should be extended to improve its behavioural realism, which is quite cumbersome, due to its complex mathematical derivation.

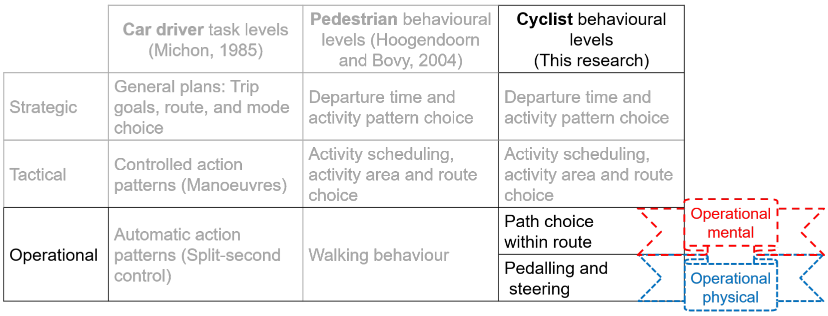
We claim that the decisions and actions taken by cyclists while riding and interacting with other traffic participants and with the infrastructure belong to the same behavioural level and should be, therefore, modelled together. We refer to this as operational cycling behaviour level and, since a proper definition of what it entails is still missing, we define it in this chapter. At the same time we put forward a novel two-layer modelling framework that can be used to capture the mental and physical processes of operational cycling behaviour. Moreover, this chapter proposes for the first time the use of discrete choice theory to identify and predict microscopic bicycle traffic flow operations. The third contribution is the application of the proposed approach to model the behaviour of cyclists when they approach a red traffic light. Discrete choice models are estimated for both layers based on trajectory data collected in Amsterdam, and describe the queue formation process, which includes selecting a queuing position and cycling towards it. The estimated models are face validated and reveal the factors that play a role in this process.

The remainder of this chapter is structured as follows. Section 2.2, then, defines the cyclist behavioural levels and describes the proposed modelling framework. In section 2.3 the proposed mathematical model is explained, followed by its application on a dataset described in section 2.4. The model estimation approach for each layer is discussed in section 2.6. In section 2.5 the results of the best performing estimated model for each layer are presented, along with simulation results for face validation. Finally, in section 2.7 conclusions are drawn and recommendations for future research are made.

2.2 Conceptual modelling framework

The focus of this thesis is on modelling operational cycling behaviour. Since literature describing the behaviour of different modes is not aligned with respect to what operational behaviour entails, the definition that will be used in this thesis needs to be provided first. Figure 2.1 shows the distinction of the behavioural levels used for car and pedestrian traffic by Michon (1985) and Hoogendoorn & Bovy (2004), respectively, and the one we propose for cycling traffic.

Twaddle (2017) adopted the definition of Michon (1985) where the operational level is limited to actions within a time horizon of milliseconds. However, according to Rasmussen (1983), riding a bicycle is a combination of tasks executed based on rules to perform manoeuvres and automatic actions for split-second control of the bicycle. We, therefore, believe that they belong to the same level, the operational level, whose time horizon is up to the order of seconds. We adopt the definitions used for pedestrians with respect to the strategic and tactical level (Hoogendoorn & Bovy 2004), whose explanation goes beyond the scope of this thesis, and focus on the operational level. The input to this level is the route from one origin to a destination. Within this level, two layers are distinguished, following the concept of a plan-action decision structure proposed by Choudhury et al. (2010) and applied to model pedestrian walking behaviour by Fukuda et al. (2014). In the upper layer, cyclists need to choose intermediate destinations and build up their path within the route while interacting with other traffic users and with the infrastructure. We call this the 'operational mental' layer. Path choices refer, among other things, to yielding, accepting a gap to merge or cross, stopping for a red traffic signal, turning, and overtaking. For the execution of each of these path choices, bicycle control dynamics in the form of pedalling and steering are necessary. This is the lower operational layer which we name 'operational physical' layer.



	Car driver task levels (Michon, 1985)	Pedestrian behavioural levels (Hoogendoorn and Bovy, 2004)	Cyclist behavioural levels (This research)	
Strategic	General plans: Trip goals, route, and mode choice	Departure time and activity pattern choice	Departure time and activity pattern choice	
Tactical	Controlled action patterns (Manoeuvres)	Activity scheduling, activity area and route choice	Activity scheduling, activity area and route choice	
Operational	Automatic action patterns (Split-second control)	Walking behaviour	Path choice within route	Operational mental
			Pedalling and steering	Operational physical

Figure 2.1: *Distinction of behavioural levels for car (left) and pedestrian traffic (middle) found in literature and for cyclist traffic (right) proposed in this thesis.*

Given this definition, we build upon the conceptual model of Gavriilidou et al. (2019b), shown in Figure 2.2, which describes cycling behaviour at the operational level, and we fit the two proposed layers within the individual behaviour, as visualised in Figure 2.3.

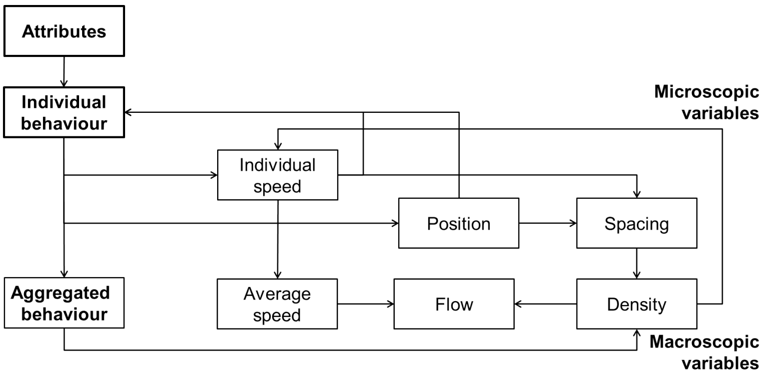
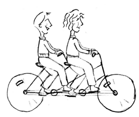


Figure 2.2: *Conceptual model of operational cycling behaviour. Attributes are linked to individual behaviour. Collectively, they lead to aggregated behaviour. These behaviours can be observed via microscopic and macroscopic variables (Gavriilidou et al. 2019b).*



Five types of path choices have been identified in the operational mental layer and each of them is captured by a separate model. The choices correspond to situations when (i) cyclists decide to overtake, but also when they are approaching an unsignalised intersection intending to cross or merge and need to decide (ii) whether to accept a gap in the conflicting traffic stream. Another choice at unsignalised intersections which is at the discretion of the cyclists is to (iii) yield to oncoming traffic. Moreover, situations at a red traffic light are covered, where cyclists decide (iv) whether they stop and (v) where to position themselves in the queue.

We hypothesise that these choices depend on a set of attributes that need to be taken into account by the models. The attributes displayed in Figure 2.3 for the decisions to overtake, yield, and stop at a red traffic light are the outcome of a stated preference survey we conducted in the Netherlands, discussed in (Gavriilidou et al. 2019b). They still need to be validated with field data, but give good insights into the behavioural attributes. The gap acceptance attributes are taken from (Jiang et al. 2013), even though they studied interactions between bicycles and motorised traffic. For bicycle-to-bicycle interactions on designated cycling infrastructure this list needs to be further investigated. In the application of the framework in this chapter, the attributes describing the queue position choice have been investigated and the findings are added to the figure.

The operational physical layer consists of the controls that each individual exerts once a path choice has been made. These controls are steering and pedalling to determine the cycling direction and speed, respectively. This layer is described by a single dedicated model covering steering and pedalling jointly. By applying these controls the state of each individual cyclist (i.e., speed, position and headway) is affected. On an aggregated scale (see Figure 2.2) they have an effect on density and other macroscopic characteristics which can, then, result in changes in the choices made by each individual cyclist, thereby substantiating an interaction between the two layers. This interaction works in two directions: (i) the choice made in the mental layer is communicated into the physical layer, and; (ii) the new state of the system after applying the decision taken in the physical layer influences the new choice to be made in the mental layer.

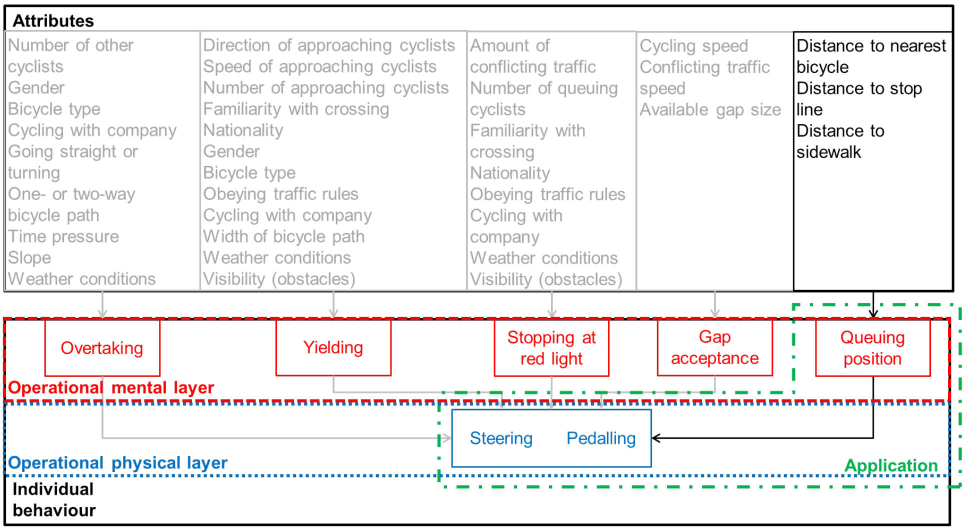


Figure 2.3: List of influential attributes per choice within the individual mental layer (coloured in red). The choice in the mental layer determines the choices within the operational physical layer (coloured in blue). The elements that go beyond the scope of this chapter are coloured in grey, while the scope of the model application in this chapter is framed within the green box (built upon Gavriilidou et al. (2019b)).

2.3 Mathematical modelling

In order to select the mathematical model that best fits our framework, we should first identify our behavioural assumptions. Those are discussed in subsection 2.3.1, followed by the description of the models for each operational behaviour layer (subsection 2.3.2 for the operational mental layer and subsection 2.3.3 for the operational physical layer).

2.3.1 Behavioural assumptions

The following assumptions are made regarding the cycling behaviour at the operational level:

1. Cyclists are effort minimisers, motivated by the general principle of least effort. This holds for both layers, though the definition of effort might differ per layer.

2. Decisions in the two layers are made sequentially.
3. The decision made in one layer is input into the other layer.
4. The updating frequency of the mental layer is smaller (i.e., has a longer horizon) than that of the physical layer.
5. When making a decision, cyclists evaluate a set of alternatives using specific attributes.

In line with the framework presented in section 2.2, we propose a two-layer mathematical model and use discrete choice theory and utility maximisation to model each layer. This allows the identification of the key attributes of the decision making process in each layer. It should be noted that discrete choice models have been used to model the movement of pedestrians (Antonini et al. 2006) and motorcycles (Lee et al. 2009; Shiomi et al. 2012). In motorcycle research, a discrete choice model has also been estimated in the context of queue formation at an intersection (Lee & Wong 2016). One of the main reasons for the lack of such a cycling model is that discrete choice models require empirical data to be estimated and there has been a lack of cyclist trajectory data which we overcome in this thesis.

2.3.2 Modelling the operational mental layer

The use of discrete choice theory to model the operational mental layer is demonstrated through its application on one of the path choices, namely the queue position (green box in Figure 2.3). The description in this subsection is qualitative, while the quantitative estimations are introduced in subsection 2.5.1.

In the operational mental layer of this application, cyclists need to decide where to stop in the queue formed upstream of a red traffic light. In our approach, the two-dimensional space is discretised in diamond-shaped cells, since we argue that they represent better the space a bicycle occupies than a rectangular grid. These cells compose the choice set. Each cell is assigned a (dis)utility based on cell attributes and characteristics of the decision maker (cyclist). Using discrete choice theory and the utility maximisation principle, a model can be estimated from cyclist trajectory data, revealing the significant attributes and their relative contribution to the overall cell utility. Availability conditions are also taken into account, since cells that are occupied cannot be re-assigned.

The diamond-shaped grid is visualised in Figure 2.4, where bicycles already present in the queue (coloured in black) are standing still and a cyclist (coloured in green) approaches and needs to make a decision. Given this situation, the green cyclist will select a cell that is not yet occupied and gives the highest utility. In this case the red cell is selected and assigned as the intended queuing position of the green cyclist. This is the output of the operational mental layer that is used by the operational physical layer.

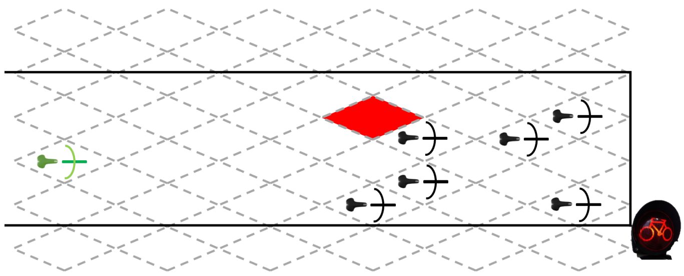


Figure 2.4: Schematic of the operational mental layer during the queuing process at a (red) traffic light. The green cyclist approaching the traffic light decides in this layer the intended queuing position (red cell) based on the characteristics of the cells, the availability conditions and utility maximisation.

2.3.3 Modelling the operational physical layer

As already mentioned, the intended queuing position is fed as input to the discrete choice model of the operational physical layer, where the cyclist decides upon the controls to reach this position. The controls are a combination of pedalling and steering, which are expressed as changes in speed and direction relative to the speed and direction, respectively, at the moment the decision is made. The justification of the choice of speed and direction difference as controls over the choice of their corresponding absolute values or the choice of a new position in the two-dimensional space in the next time step relates to the assumption that cyclists are effort minimisers. This means that they choose the relative effort they are willing to exert in each time step and that goes through changes in pedalling and steering rather than anticipation of their future position.

The choice alternatives are visualised in the fan-shaped individual-specific grid in Figure 2.5. The fan shape is selected because it reflects the angular movements that characterise cyclist motion. The angular sections capture the radial directions accessible with appropriate changes in steering. The number of angular sections and arched zones is only illustrative and should be determined dependent on the application. In this example, the middle angular section corresponds to no change in the direction, two sections to the right correspond to a small and a bigger steering movement towards the right, and sections to the left steering to the left. The arched zones represent possible relative speed regimes that can be reached through pedalling or braking. The arch closest to the bicycle corresponds to speed reduction (deceleration), the arch furthest away corresponds to an increase in speed (acceleration) and the middle arch corresponds to a choice of no change in speed.

Two more aspects are demonstrated in the figure. One is the sequence of decisions in time within this layer (a lighter shade of grey is given to the grid for decisions to be made in each future time step). The sequence of positions resulting from these choices leads the cyclist to the final position and together forms the cyclist trajectory. The other aspect demonstrated is that the grid is always aligned with the cycling direction of the cyclist at the moment the decision is made. This is shown by the rotation of the grid in each time step, such that the 'no change in direction' alternative is a continuation of the change in direction that was chosen in the previous time step. The centre of the grid in each time step corresponds to the new location of the cyclist, which depends on the time step, the cycling speed and the choice of change in speed made in the previous step. As the figure is illustrative and no numerical values for speed and time are assigned, the distance separating the different grids is only qualitative.

At each time step, the cell in the fan with the highest utility is selected and the position of the cyclist is updated for one time step, when a new decision is required. It is possible that in between these decision moments there is an interaction with the operational mental layer if a situation occurs that was originally not anticipated by the cyclist and necessitates the estimation of a new intended queuing position (e.g., if another cyclist occupies the originally intended position).

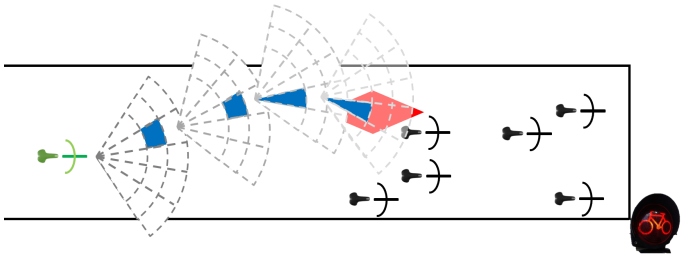


Figure 2.5: Schematic of the operational physical layer during the queuing process at a traffic light given the intended queuing position (red cell) provided by the operational mental layer. A sequence of decisions (blue cells) is made that corresponds to the combination of angle and speed difference with the highest utility at each time step (grey-scale fans). The sequence of positions resulting from these choices forms the cyclist trajectory.

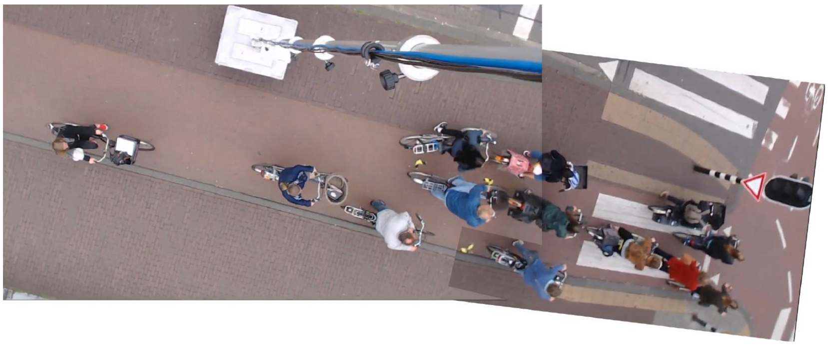
In any case, the result of this interaction between the two layers and the decisions made over time is the cyclist trajectory to reach the intended queuing position. This trajectory together with the final queuing position fully describe the operational cycling behaviour. A quantitative application of the proposed framework is presented in section 2.5, where specific models have been estimated using field data from Amsterdam, the Netherlands. Prior to that, the data available for the model identification is described in section 2.4.

2.4 Data on queue formation process

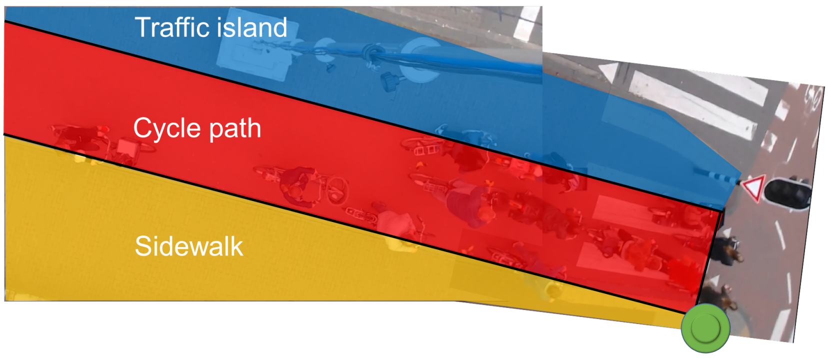
This section presents the dataset used for the model estimation and validation. First, both the site and dataset are introduced (subsection 2.4.1), followed by a description of the data processing to prepare the dataset to estimate the models (subsection 2.4.2).

2.4.1 Site and dataset description

The dataset used for the model estimation contains cyclist trajectories that have been extracted from video camera footage at a signalised intersection in Amsterdam, the Netherlands (Figure 2.6(a)). The site consists of a 2m-wide unidirectional cycle path that is separated on the left from motorised traffic (even though the cycle path itself can be used by scooters) via a traffic island that has a different surface type but no height difference and can be used by pedestrians and cyclists who queue. The cycle path is also separated from the sidewalk by a curb on the right. These areas are illustrated in Figure 2.6(b). With two cameras elevated over the cycle path, top-view video images were recorded during an afternoon. The combined view of both cameras covers a length of 20m upstream of the stop line at the traffic light.



(a) Combined view from front and back camera at the intersection. The yellow tape on the cycle path marks the overlapping area between the two cameras.



(b) Study areas of the traffic island, the cycle path and the sidewalk. The black lines denote the edges of the cycle path and a green button is used to show the location of the 'request-green' button.

Figure 2.6: Top view and areas of interest at the site.

As the aim of our application is to use the two-layer framework to model the queue formation process (i.e., choosing a location to queue and cycling towards it), trajectories of cyclists approaching the traffic light during the red-light phase are collected. We focus on pure bicycle-to-bicycle interactions and therefore, only bicycle trajectories are extracted. Red-light phases where more than one scooter was present are omitted, as well as phases when there is interaction with crossing pedestrians. The phases with one scooter are kept, assuming that one scooter does not have an effect on the results when considering pure bicycle-bicycle interactions. Only the position it occupies is tracked, so that it is not available to cyclists that arrive later. This way the sample of tracked bicycles increases, and the trajectory of the scooter is ignored.

The last criterion for a red-light phase to be removed from the dataset is the presence of disturbances, such as pedestrians crossing and creating conflicts, bicycles joining the queue from another side or even the sidewalk where they were parked, and cyclists that decided to run the red light even though they had originally queued, thereby initiating movements within the queue.

The final dataset consists of 46 red-light phases with 454 cyclists and 18 scooters in total queuing up. It should be noted that cyclists arriving after the traffic light turned green were not included since their intended queuing position, if any, was not observed. The transition from video files to microscopic cyclist trajectories comprises six steps, which were performed as follows:

1. Decomposition of videos into frames with an average frame rate of 6 frames per second (fps).

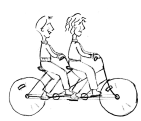
2. Manual tracking per frame of the head of each cyclist who approaches the intersection during a red-light phase until standstill.

3. Height transformation to project the trajectories at the head positions to the ground.

4. Orthorectification to correct for the distortion due to the fact that the cameras were placed at an angle and did not point vertically downwards to the cycle path, as well as to compensate for the lens distortion.

5. Time conversion from frame number to seconds.

6. Trajectory merging of the two cameras for each cyclist.



For more information about the data collection and the extraction steps, the reader is directed to the paper by Goñi Ros et al. (2018).

2.4.2 Data processing

Figure 2.7 shows the cyclist trajectories, speeds and steering angles of one red-light phase. The raw trajectories have been extracted by means of manual clicking on the heads of the cyclists and thus contain noise ('*' in Figure 2.7). To remove this noise, a smoothing process is applied on the data. The smoothing is done by means of a moving average with a fixed-length sliding window across the trajectory data vectors (x: vector of positions in the horizontal direction, y: vector of positions in the vertical direction, t: vector of time instances corresponding to each position). The calculation of the mean is performed for each element of the original vector while centring the window around the corresponding element. Different sliding window lengths were compared (see appendix A). The results only have limited difference and favour the smoothing of the trajectories over a duration of 6 frames. The smoothed data points are marked by '+' in Figure 2.7.

Although the average frame rate was 6fps, it was not constant over time, resulting in data points that are not separated by the same time gap. For our application, it is crucial to have a consistent time discretisation throughout all cyclist trajectories as each point corresponds to a moment at which a decision was made. These points are derived by taking the instant right before these homogenised timestamps and projecting the trajectory in (x,y) using the smoothed speed at that instant. Different time steps were compared (see appendix A) and a time step of 1 second was found to be best at muting the noise. This means that each cyclist is assumed to make a new decision regarding the steering angle and speed difference every second. The final data points are marked by 'o' in Figure 2.7.

The trajectories show the path each cyclist followed from the moment they were detected by the back camera up to their final queuing position. The values in both axes have been adjusted for the visualisation, such that the (0,0) point coincides with the location of the 'request-green' button, while in the actual dataset they have positive values that increase in the direction of cyclist movement. The speed and steering angle are computed between consecutive data points and are visualised relative to the horizontal distance that the cyclist has traversed. The horizontal axis of these two graphs has been offset such that all trajectories end at the same point, which facilitates the comparison between the original, the smoothed and the final dataset. A number of observations can be made:

- The trajectories show a good match between the original, the smoothed and the final dataset.
- The smoothing helps reducing the noise in the speed and steering angle.
- The speed is decreasing throughout the observed trajectory which is in line with the findings of Jiang et al. (2013) that deceleration occurs within the 30 m upstream of the intersection.
- The steering adjustments are small at the beginning of the trajectory.
- When the speed is low, the steering adjustments increase and they are maximum at the end of the trajectories where the speed is the lowest and the head is swaying more.

These final trajectory points are used for the model estimation. The operational mental layer requires only the last point which corresponds to the queuing position of each cyclist. The operational physical layer takes into account every point as they have been assumed to correspond to a decision point. It should be noted that the reason why the operational mental layer does not make use of the original dataset is to guarantee the consistency between the two layers.

2.5 Model estimation approach

In this section the estimation approach for each layer (subsection 2.5.1 for the operational mental layer and subsection 2.5.2 for the operational physical layer) is discussed. It includes a justification of the grid choice and selection of attributes to explain the corresponding behaviour, as well as assumptions specific to the model estimation.

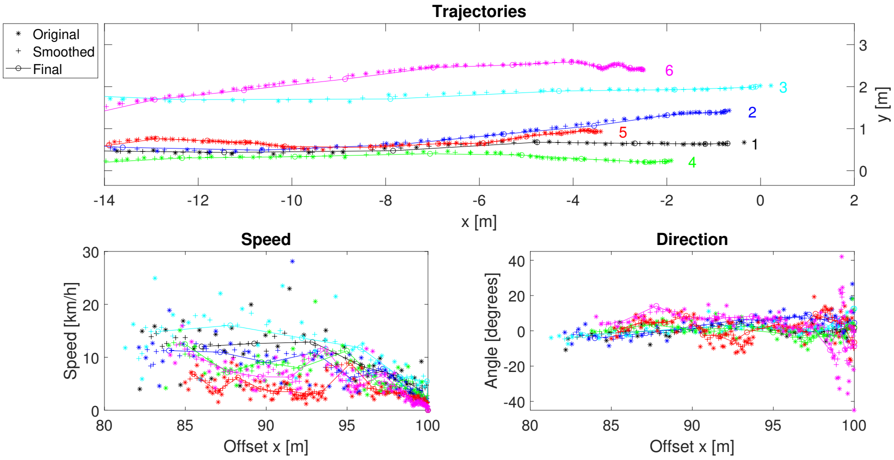


Figure 2.7: Cyclist trajectories numbered based on their order of arrival (top), speed (bottom left) and steering angle (bottom right) when approaching a red traffic light in the original dataset, when smoothed with a sliding window length of 6 frames and finally when homogenised with time step of 1s. In the top figure the point (0,0) is the location where the stop line meets the curb of the sidewalk, while at the bottom figures the positions in x have been offset such that they end at the same arbitrary point for all cyclists.

2.5.1 Operational mental layer estimation approach

This layer aims to capture the decision making when joining a queue, where cyclists have to choose their queuing position, as introduced in Figure 2.4. The observed queuing positions (i.e., last trajectory point of the processed dataset when cyclists are at standstill) are visualised in Figure 2.8, where the point (0,0) is the location of the ‘request-green’ button and the red lines indicate the boundaries of the cycle path. As expected, positions next to the ‘request-green’ button are the most frequently selected. Other positions at the stop line are also favourable, as well as positions on the traffic island, especially for cyclists who want to make a left turn at the intersection. As the queue increases in length, there is a preference for a position next to the curb of the sidewalk rather than a position in the middle of the cycle path.

These choices can be analysed to identify which attributes have an influence, and to what extent, on the queuing position choice by estimating a choice model. The estimation requires the definition of the choice set, the specification of the utility functions and the demarcation of the estimation assumptions.

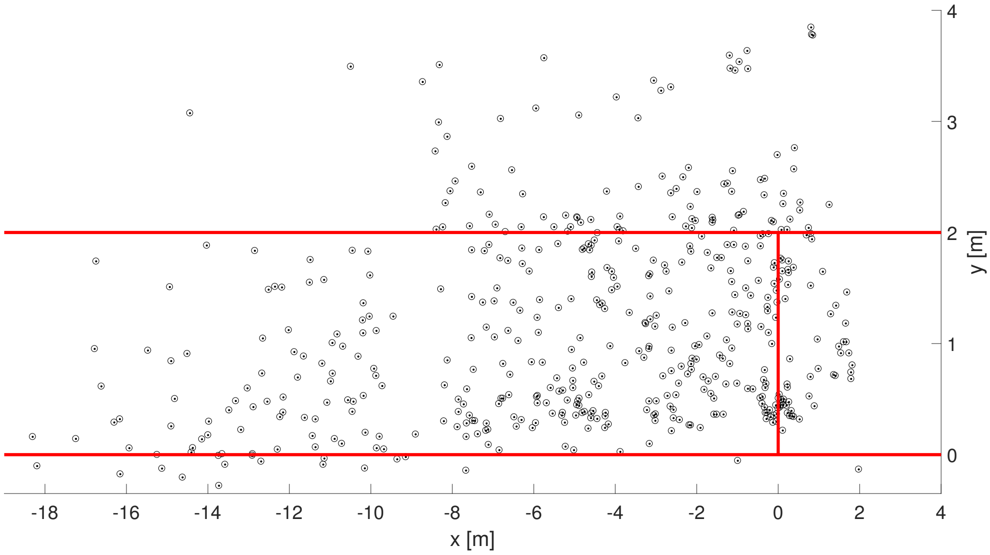


Figure 2.8: Observed queuing positions. The point (0,0) is the location where the stop line meets the curb of the sidewalk (also the location of the 'request-green' button). The red lines indicate the boundaries of the cycle path.

Choice set definition

The cycle path and surrounding areas (sidewalk and traffic island next to the cycle path) are discretised using the aforementioned diamond-shaped grid. The resulting cells correspond to the choice alternatives of a cyclist. As previously mentioned, this grid better captures the shape of the bicycle (compared to a rectangular grid) and allows for a more realistic representation of queuing.

Each cell is scaled such that it can fit one cyclist based on the standard dimensions of 2 m length and 70 cm handlebar width (CROW 2016). These dimensions also show a good match with the average observed spacing in the longitudinal and lateral direction between stopped cyclists in our dataset.

The grid is generated in such a way that there is a cell right next to the sidewalk whose middle crosses the stop line. This is because the 'request-green' button is located right next to the stop line so cyclists stop on top of the line rather than behind it.

The real queuing position is then projected to this grid and assigned to the cell whose centroid is closest to it. This projection results in the top plot of Figure 2.10, which will be explained in the subsection 2.6.1. The choice set for each cyclist comprises the cells that are not already occupied by others.

Utility specification

The attributes (cell characteristics) that are hypothesised to capture the attractiveness (utility) of a cell are the distance to the stop line, the distance to the edges of the cycle path and the presence of other cyclists in the queue. Moreover, the behaviour of the first cyclist is hypothesised to be different from the behaviour of the rest, since the first arriving cyclist needs to stop next to the 'request-green' button to be able to press the button to request green. For this reason, the attributes related to the distance to the stop line and the distance to the edges of the cycle path are separately estimated for the first cyclist and for the rest. Given these hypotheses, a general description of the specific attributes is first provided, followed by a full list of the detailed attribute notation and definition.

Regarding the distance to the stop line, cells whose centroid is downstream the stop line are differentiated from those that are upstream. This way the former, i.e. stopping after having crossed the stop line, can be penalised and avoided as a queuing position.

With respect to the distance to the edges of the cycle path, it is hypothesised that the effect on utility is not symmetrical as the distance increases within and outside of the two edges, because being on the cycle path is desirable, while being on the sidewalk is less comfortable due to the presence of pedestrians and being on the traffic island increases the proximity to motorised traffic. Based on this hypothesis, the area covered in the choice set is subdivided into four sublanes, namely the sidewalk, the right lane of the cycle path, the left lane and the traffic island. The right edge of the cycle path is taken as reference for the first two sublanes and the left edge as reference for the last two.

The presence of other cyclists can be represented in several ways. Therefore, more than one attribute is defined. One way is the distance to the nearest cyclist in the queue. Another considers the number of cyclists within each of the aforementioned sublanes, as the more cyclists stopped within a sublane, the less attractive the sublane becomes, because cyclists cannot manoeuvre to overtake and would need to join the end of the queue. This end of the queue may also be seen as an offset of the stop line within each sublane, i.e. cyclists have a higher utility in stopping closer to the end of the queue. An attribute is therefore added that considers the distance to the cyclist at the back of the queue of the sublane.

The full list of attributes for this layer is given in Table 2.1.

In the construction of the systematic part of the utility functions (V), interaction terms among these attributes are used in a linear weighted summation. The weights (coefficients) are denoted by β and are generic for all alternatives as there is no straightforward way to classify them in nests that would acquire alternative specific weights. An example utility function of a cell c (V_c) in a model where only the interaction term between the dummy X_F, the dummy X_up, and the variable X_d2stop is considered, is shown in Equation 2.1.

$$
\begin{aligned}
V_\mathrm{c} = & \\
& \beta_\mathrm{upF} \cdot X_\mathrm{F} \cdot X_{\mathrm{d2stop_c}} \cdot X_{\mathrm{up_c}} + \beta_\mathrm{upR} \cdot (1 - X_\mathrm{F}) \cdot X_{\mathrm{d2stop_c}} \cdot X_{\mathrm{up_c}} \\
& + \beta_\mathrm{downF} \cdot X_\mathrm{F} \cdot X_{\mathrm{d2stop_c}} \cdot (1 - X_{\mathrm{up_c}}) \\
& + \beta_\mathrm{downR} \cdot (1 - X_\mathrm{F}) \cdot X_{\mathrm{d2stop_c}} \cdot (1 - X_{\mathrm{up_c}})
\end{aligned}
\tag{2.1}
$$

Estimation assumptions

When estimating a model within this layer, the following assumptions are made to simplify the estimation process:

1. There is no correlation between the alternatives (Independence of Irrelevant Alternatives, IIA property) and therefore, a multinomial logit model can be used.

Table 2.1: List of variables for the operational mental layer.

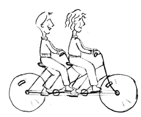
Attribute	**Unit**	**Explanation**
X_{F}	-	dummy indicating whether the cyclist is the first one arriving
X_{button}	-	dummy to denote if a cell is the cell next to the ‘request-green’ button
X_{d2stop}	m	longitudinal distance between the location of the stop line and the centroid of the considered cell
X_{up}	-	dummy indicating whether the considered cell is upstream or on the stop line
X_{onside}	-	dummy to denote if a cell is on the sidewalk
X_{onisland}	-	magnitude of steering intensity when considering changes towards the left
X_{rightln}	-	dummy to denote if a cell is on the right lane of the cycle path
X_{leftln}	-	dummy to denote if a cell is on the left lane of the cycle path
X_{d2Redge}	m	absolute lateral distance between the location of the right edge of the cycle path and the centroid of the considered cell
X_{d2Ledge}	m	absolute lateral distance between the location of the left edge of the cycle path and the centroid of the considered cell
$X_{\mathrm{d2nearEucl}}$	m	Euclidean distance between the centroid of the considered cell and the one closest to it that is occupied by cyclists already standing in the queue
X_{d2nearX}	m	minimum absolute longitudinal distance between the location of the centroid of the considered cell and those already occupied by cyclists standing in the queue
X_{d2nearY}	m	minimum absolute lateral distance between the location of the centroid of the considered cell and those already occupied by cyclists standing in the queue
X_{total}	cyclists	total number of cyclists within the sublane of the considered cell
X_{d2lastX}	m	longitudinal distance between the last queuing cyclist in the sublane where the considered cell belongs and the centroid of the considered cell, if the considered cell is upstream

2. Queuing spots are assigned upon the entrance of a cyclist in the camera vision field, which means that this decision precedes any decisions on the operational physical layer and that the assignment of spots follows the order of arrival of cyclists.

3. The assigned queuing spots are not updated over the course of the cyclist trajectory. This means that the interaction between the two layers is in this application one-way.

4. Cyclists are assumed to be aware of the spot selected by their predecessors. This is imposed through availability conditions in the logit model, which remove cells that are already assigned to a cyclist from the choice set of oncoming cyclists.

These assumptions can later be relaxed, for example by considering spatial correlation between the diamond-shaped cells and by allowing the updating of the decision for the intended queuing position. Part of the spatial correlation has already been captured by the attributes related to the distance from the stop line and the edge of the cycle path. But as cyclists do not see the diamonds on the cycle path, they might apply a different discretisation of space in areas combining several cells. This should be considered in future research and would require the definition of areas that are correlated and the estimation of more advanced logit models. However, a first attempt that considered a nesting structure around the 'request-green' button did not show evidence of such spatial correlation. The decision updating becomes relevant when unanticipated changes take place, such as a cyclist entering the cycle path from another direction and occupying the originally desired position. Another reason to consider updating is when speed differences are large and overtaking might place. In this case, the first come first serve rule might need to be replaced by a rule based on cycling speed. Since the dataset does not contain disturbances of sudden appearing cyclists and there is no information on desired queuing positions other than the revealed one, these effects are righteously ignored.



2.5.2 Operational physical layer estimation approach

Within this layer, the intended queuing position is known and the cyclist decides in every time step the changes in pedalling and steering until the next time step. In order to reach the intended queuing position, a sequence of time steps, and corresponding decisions, is needed and results in the cyclist trajectory.

The estimation requires the definition of the choice set, the specification of the utility functions and the demarcation of the estimation assumptions.

Choice set definition

In the physical layer, cyclists decide whether they will change their speed and direction at the current time step. By looking at the combinations that occur in the processed dataset the discretisation of the fan-shaped grid of Figure 2.5 can be motivated.

The observed choices are visualised by the blue dots in Figure 2.9. It shows that most observations are concentrated around zero in the angle difference and more specifically in the boundary of [-15,15] degrees. Larger changes in the steering angle only take place when there is no, or very small, change in the speed (ΔSpeed between [-2,2] km/h). From further inspection of the dataset, it is noted that this coincides with very low speeds. Moreover, regarding the speed changes, most observations are negative, which is in line with the findings of section 2.4 that cyclists are already decelerating when entering the observed area.

Following these insights, for our application the fan-shaped choice grid is defined to range from speed changes of -12 km/h to +8 km/h with a step of 2 km/h, and the steering angle changes in degrees are {-45, -30,-15,-10,-5,0,5,10,15,30,45}. The observed choices are then assigned to their closest grid point and choices that would result in a negative cycling speed are made unavailable.

Utility specification

Based on the observation that the distance to other cyclists and the curb play a role in the decision for the queuing position, we hypothesise that they also affect the path that is chosen to reach that position. In this decision layer, the position of the stop line is less relevant but what might have an effect is the distance to the intended queuing position. When considering this distance, one can differentiate between choices that result in passing the intended queuing position and choices that do not. This way the former can be penalised and avoided towards the end of the trajectory. Additionally, it is hypothesised that the behaviour towards moving cyclists differs from the behaviour against stopped cyclists or obstacles. This discrepancy is captured by the distance to the nearest bicycle and the difference in cycling speed, which is calculated separately for moving and for stationary bicycles. The full list of attributes is given in Table 2.2.

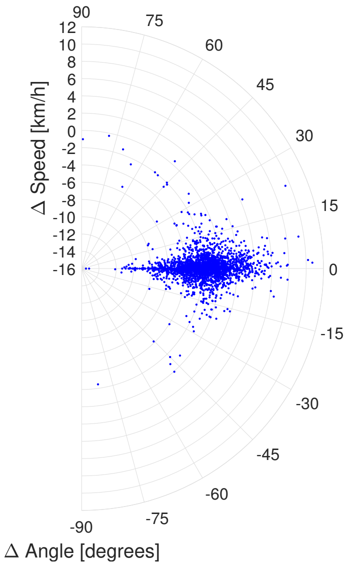


Figure 2.9: Observed choices of changes in steering angle and speed.

These attributes are used in a linear weighted summation to construct the systematic part of the utility functions (V), with the exception of the distance to the destination, which is covered by an interaction term between the dummy X_{pass} and the X_{d2dest}. An example utility function of alternative a (V_{a}) in a model where only this interaction term is considered, is shown in Equation 2.2.

$$V_{\text{a}} = \beta_{\text{over}} \cdot X_{\text{d2dest}_{\text{a}}} \cdot X_{\text{pass}_{\text{a}}} + \beta_{\text{under}} \cdot X_{\text{d2dest}_{\text{a}}} \cdot (1 - X_{\text{pass}_{\text{a}}}) \tag{2.2}$$

Table 2.2: List of variables for the operational physical layer.

Attribute	**Unit**	**Explanation**
X_{d2dest}	m	Euclidean distance between the destination (i.e., the intended queuing position) and the location to be reached within a time step if the considered change in speed and angle is chosen position
X_{pass}	-	dummy indicating whether in the next step, given the considered change in speed and angle, the cyclist will have traversed a longer longitudinal distance than needed to reach the intended queuing position
X_{d2Mov}	m	minimum Euclidean distance between the anticipated positions of the cyclists in front who are moving, and the location to be reached within a time step if the considered change in speed and angle is chosen
X_{d2Stop}	m	minimum Euclidean distance between the anticipated positions of the cyclists in front who are stopped, and the location to be reached within a time step if the considered change in speed and angle is chosen
X_{spdMov}	m/s	maximum speed difference between the considered speed and that of cyclists in front who are moving
X_{spdStop}	m/s	speed to be reached if the considered change in speed is chosen and if there are cyclists in front who are stopped. Since the stopped cyclists have no speed, this attribute reflects the disutility of having the considered speed when others have stopped
X_{dVmov}	m/s	maximum speed difference between considered speed and speed of moving cyclists in front
X_{step}	-	dummy indicating whether in the next step the cyclist will need to get on or off the curb of the sidewalk
X_{offpath}	-	dummy indicating whether in the next step the cyclist will need to get on or off a the traffic island

Estimation assumptions

When estimating a model within this layer, we make the following assumptions:

- Changes in cycling speed and direction are decided simultaneously.
- Cyclists do not move backwards and so no negative speeds are allowed.
- Decisions made by the same person at different time steps are independent. As no serial correlation is assumed, a multinomial logit model is to be estimated.
- When other cyclists are present, there is full knowledge of their current speed and direction.
- Zero acceleration is assumed for the other cyclists and so their anticipated position within one time step can be estimated based on their current cycling speeds.
- There is no memory from previous time steps. Only cyclists within the vision field at the current position affect the decision to be made.
- The vision field contains everything that is in front of or at least at the same longitudinal position as the cyclist making a decision.

The pitfall of ignoring the serial correlation is that bias of an individual towards a certain type of behaviour is overlooked, and the risk of inconsistent behaviour between time steps is introduced. Even though it decreases the model realism, we argue that it is an acceptable simplification to get first insights into the operational physical layer. If this assumption is relaxed and panel data are considered, a mixed logit model should be estimated. In terms of anticipation, the assumption of zero acceleration is reasonable as it cannot be expected that the intentions of others are known. The full knowledge of the speed and position assumption could be relaxed by considering some noise rather than the exact measurements. Last but not least, future research should introduce a memory function to improve the anticipatory skills of cyclists and increase the model realism.

2.6 Results and discussion

This section provides and discusses the model estimation results for each layer. Models have been estimated using Python Biogeme (Bierlaire 2016). The best performing model is found based on goodness of fit measures, i.e. $\bar{\rho}^2$, AIC and BIC criteria, and is presented in subsection 2.6.1 for the operational mental layer and in subsection 2.6.2 for the operational physical layer. Both models are face validated by means of simulation with Biogeme. The simulation results are discussed in subsection 2.6.3.

2.6.1 Operational mental layer model

The estimated values of the coefficients of the best performing model are shown in Table 2.3, along with their robust statistics. The model consists of 12 parameters, one third of which captures the behaviour of the first arriving cyclist. These four attributes are differently weighed from the corresponding ones for the rest of the cyclists, which confirms the hypothesis made in subsection 2.5.1 that the behaviour of the first cyclist is different.

The cell next to the 'request-green' button has a positive coefficient ($\beta_{\text{buttonF}} = 1.24$) for the first cyclist, while it does not affect the utility for the rest of the cyclists. Moreover, the utility decreases the further the queuing position is from the stop line. This disutility is greater in the case of crossing the stop line and stopping further downstream ($\beta_{\text{downF}} = -2.13$) compared to stopping upstream of the stop line ($\beta_{\text{upF}} = -1.18$). Regarding the distance to the edges, as most cyclists stop within the cycle path, the only attribute with sufficient observations to estimate a coefficient is the distance from the curb of the sidewalk to a position within the right lane of the cycle path. This coefficient has a positive value ($\beta_{\text{rightlnF}} = 4.91$), which shows that first arriving cyclists prefer to be close to the middle of the cycle path (i.e. at the end of the right lane). This is reasonable since it serves the purpose of stopping next to the 'request-green' button.

For the rest of the cyclists, three coefficients are estimated concerning the distance to the edges. There is an increase in utility by being on the right lane ($\beta_{\text{rightlnR}} = 1.21$), and a decrease by being on the sidewalk ($\beta_{\text{onsideR}} = -6.46$) and on the traffic island ($\beta_{\text{onislandR}} = -1.85$).

The difference in magnitude of the inflicted disutility can be explained by the fact that the sidewalk is primarily intended for use by pedestrians, while the traffic island can be used by cyclists, especially if they want to turn left at the intersection. Also for these cyclists, there is a disutility the further downstream the queuing position is from the stop line ($\beta_{\text{downR}} = -1.29$). The coefficient of the attribute describing the distance to the stop line for cells upstream the stop line is positive ($\beta_{\text{upR}} = 0.30$), which might seem counter-intuitive, but can be explained by its negative correlation with the distance to the nearest bicycle, as well as with the distance to the last stopped cyclist within a sublane. These two have negative coefficients $\beta_{\text{d2nearX}} = -0.53$ and $\beta_{\text{d2lastX}} = -0.22$, respectively, which indicates that cyclists want to stay close to each other in the queue in the longitudinal direction. Since the last stopped cyclist within a sublane is considered as an offset of the stop line, it is reasonable that the coefficient is negative and the arriving cyclist wants to stay as close as possible to the adjusted stop line. The last attribute of the model captures the effect of the number of queuing cyclists within a sublane and, as expected, has a negative coefficient ($\beta_{\text{total}} = -0.39$). This shows that the more cyclists stopped within a sublane, the lower the utility of that sublane and therefore, it is more likely that the arriving cyclist will choose to stop in another sublane.

2.6.2 Operational physical layer model

The estimated values of the coefficients of the best performing model are provided in Table 2.4, along with their robust statistics. All values are statistically significant, which confirms that the hypothesised attributes influence the choices made with respect to changes in speed and cycling direction.

Moreover, the coefficient values prove that the behaviour towards stopped and moving cyclists is indeed different, especially when considering the distance. There is a higher disutility when getting closer to a stopped cyclist than to a moving one. This can be explained by the fact that the moving cyclist continues to change position, while stopped cyclists form a (static) obstacle when the intended queuing position is not adjacent to them.

Table 2.3: Estimated model parameters for the operational mental layer.

Coefficient name	Coefficient value	Robust standard error	Robust t-test	Robust p-value
β_{buttonF}	1.24	0.37	3.31	0.00
β_{upF}	-1.18	0.26	-4.58	0.00
β_{downF}	-2.13	0.55	-3.90	0.00
β_{rightlnF}	4.91	0.75	6.57	0.00
β_{upR}	0.30	0.03	9.70	0.00
β_{downR}	-1.29	0.31	-4.14	0.00
β_{rightlnR}	1.21	0.21	5.89	0.00
$\beta_{\text{onislandR}}$	-1.85	0.23	-8.05	0.00
β_{onsideR}	-6.46	1.04	-6.19	0.00
β_{d2nearX}	-0.53	0.05	-10.90	0.00
β_{total}	-0.39	0.06	-6.43	0.00
β_{d2lastX}	-0.22	0.04	-6.30	0.00

With respect to the distance to the intended queuing position, there is a penalty for changes in speed and angle that increase this distance. The penalty is bigger when the position is passed, which is reasonable. Cyclists should not be willing to cycle further than their intended queuing position.

Another valuable insight is the disutility of having to go on and off the cycle path at consecutive time steps. As expected, the disutility is much higher on the side of the sidewalk due to the presence of the curb, while on the side of the traffic island the surfaces are on the same level and only the surface type changes.

A first attempt to consider the panel data in a mixed logit model showed evidence of correlations present in the choice for no changes in the steering angle, but did not significantly affect the coefficient values of the simple multinomial logit model. Further research is needed to investigate the effect of the panel data in more detail.

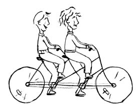


Table 2.4: Estimated model parameters for the operational physical layer.

Coefficient name	Coefficient value	Robust standard error	Robust t-test	Robust p-value
β_{under}	-1.01	0.05	-19.61	0.00
β_{over}	-2.04	0.10	-20.75	0.00
β_{d2Mov}	-0.40	0.04	-9.06	0.00
β_{d2Stop}	-0.24	0.05	-5.24	0.00
β_{spdMov}	-0.93	0.06	-15.73	0.00
β_{spdStop}	-0.61	0.06	-10.77	0.00
β_{step}	-2.46	0.20	-12.44	0.00
β_{offpath}	-0.83	0.09	-9.08	0.00

2.6.3 Face validation using simulation

Using the estimated parameters for each model, a simulation is performed, where a prediction is made for each observation in the dataset. In this case, we have used the same dataset for estimation and face validation, as we do not have another dataset available and the dataset is too small to segment. Having the same dataset means that the attributes and availability conditions describing the situation at which every individual made a decision remain the same. The simulation uses the estimated model to compute all utility functions and the probabilities of each alternative. The individual's probabilities of an alternative are aggregated by averaging over all individuals to whom the corresponding alternative was available. The true (observed) choices can then be compared with the predicted (simulated) ones.

The comparison for the operational mental layer is visualised in Figure 2.10, where the white dot at point (0,0) is the location of the 'request-green' button and the red lines indicate the boundaries of the cycle path. The observed choices show a preference for the right lane of the cycle path, which is well reproduced in the simulation results. Moreover, the choice of the first arriving cyclists to stop next to the 'request-green' button is very accurately replicated by the simulation. Another observation is that cyclists at the front of the queue are likely to stop at any lateral position on the cycle path or on the traffic island, while the longer the queue grows, the most preferable the right lane becomes, possibly due to the presence of the curb so the cyclists can rest their foot. This trend is also captured by the model; the front cells on the left lane have a higher probability than those upstream and the probability decreases with the longitudinal distance. These simulation results are, therefore, considered a good representation of reality.

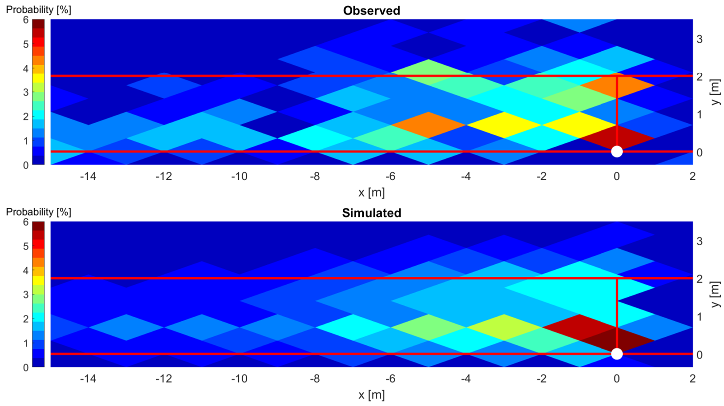


Figure 2.10: Probability of a diamond cell being selected as the queuing position in the observed (top) and the simulated (bottom) dataset. The white dot, point (0,0), is the location where the stop line meets the curb of the sidewalk (also the location of the 'request-green' button). The red lines indicate the boundaries of the cycle path.

The comparison for the operational physical layer is visualised in Figure 2.11. The pattern displayed in the two fans is similar with observed choices having less variance and thus higher probability values for no change in direction and slight deceleration, while the simulated choices are more scattered. The trend of speed reductions and small changes in the steering angle is captured well by the simulation. In order to present these results more quantitatively in a single assessment value, the positions in x and y that result from the choices of speed and angle change are calculated. The absolute percentage error made in each observation i can be computed per direction (i.e., x and y) using the formulas

$$
\begin{aligned}
x_{\mathrm{error_i}} &= \frac{|x_{\mathrm{sim_i}} - x_{\mathrm{obs_i}}|}{x_{\mathrm{obs_i}}} \\
y_{\mathrm{error_i}} &= \frac{|y_{\mathrm{sim_i}} - y_{\mathrm{obs_i}}|}{y_{\mathrm{obs_i}}}
\end{aligned}
\tag{2.3}
$$

The mean absolute percentage error (MAPE) in the longitudinal x direction is 4.37% and in the lateral y direction 1.79%. These values are considered low and prove that the model generates plausible results.

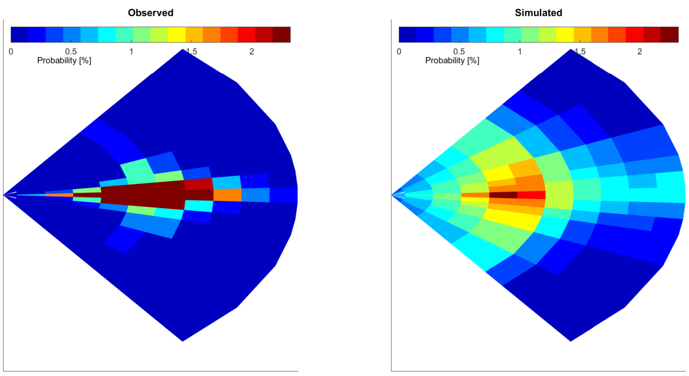


Figure 2.11: Probability of a combination of change in steering angle and speed to be selected in the observed (left) and the simulated (right) dataset.

2.7 Conclusions and recommendations

In this chapter, different behavioural levels for cyclists have been defined for the first time, while focusing on the operational level. We hypothesised that this level consists of two intertwined processes, namely the path choices made within a route and the bicycle control dynamics through changes in pedalling and steering. We put forward a novel two-layer framework to capture the tasks within the mental and physical layers of the operational level. Discrete choice theory was proposed to model each layer and the plausibility of the framework was demonstrated through an application. Using cyclist trajectory data from a signalised intersection in Amsterdam, the Netherlands, models were estimated and face validated. The models describe the behaviour of cyclists when approaching and queuing at a red traffic light, including selecting a queuing position (operational mental layer) and cycling towards it (operational physical layer). The models reveal the attributes that influence queuing behaviour.

For this specific application of the modelling framework, we found that when deciding on a queuing position, the first arriving cyclist behaves differently than the rest as there is the need to stop next to the ‘request-green’ button and press it. Additionally, cyclists prefer to stop on the right lane of the cycle path and upstream of the stop line. Positions on the sidewalk are less favourable than those on the traffic island, because the former are hindered by pedestrians on the sidewalk, while the latter can be attractive for left-turning cyclists. Furthermore, cyclists prefer to stop close to each other, but once a sublane becomes crowded, they prefer to go to another sublane. This disutility is traded off with their desire to stay on the right lane and once the front stopping positions are occupied, there is a trend to stop closer to the curb of the sidewalk rather than build up all sublanes equally. These results are intuitive because cyclists want to use the curb as a resting position when stopped and as an assist when accelerating. Once this intended queuing position is decided upon, the cyclists need to create a trajectory towards it, which they do through changes in their speed and steering angle at regular time intervals with the aim of reaching that position. Based on our estimation results, cyclists behave differently towards stopped and moving cyclists, which is reasonable since stopped cyclists form an obstacle on the way and an increase of the speed difference might lead to unsafe situations that are preferably avoided. Moreover, they are attracted by their intended queuing position and deter from passing it. They additionally deter from changing surface type and, even more strongly, from stepping on and off curbs.

These findings provide valuable insights for the design of cycling infrastructure. One way to avoid long sparse queues would be to provide an elevated curb on both sides so that cyclists can use it as a resting spot. This elevation is also advisable to prevent cyclists from leaving the cycle path and interfering with pedestrian traffic. When it is not possible, changing the surface type can be an alternative measure. The reason why sparse queues should be avoided is the fact that dense queues have shorter discharge times (Goñi Ros et al. 2018), so their green phase and the cycle time of the intersection can be reduced.

The simulation results reproduce patterns observed in the empirical data and thereby demonstrate the face validity of the models. However, there is room for improvement, which could be sought in including other attributes, such as the time of day or weather conditions, or adding personal characteristics such as age, gender, bicycle type and the riding direction after the light is green. Moreover, heterogeneity between cyclists could be considered by drawing the coefficients from a distribution rather than fixing them to one value for everyone. Furthermore, the assumption with respect to the independence of alternatives could be lifted and models that allow for correlation of alternatives, such as cross-nested or mixed-logit, could be estimated. Other modelling assumptions could be tested in future research as well, such as the vision field and the cyclists in it that are taken into account, or the anticipation and memory skills of the cyclists.

Apart from improving the currently estimated models, a future research direction is their adjustment to enable the communication of the two layers and potentially updating the intended queuing position decision. Additionally, the models could be validated on other intersections and extended for other datasets where scooters and pedestrians are present so that their effect is captured as well.

Last but not least, the generalised application of the proposed approach should be substantiated by estimating models for other choice situations in the conceptual framework. These models can then be used in microsimulations and to update the model attributes in the framework. A challenge in this process has been the shortage of cyclist trajectory data which we will tackle in future work thanks to the dataset collected through our controlled large-scale cycling experiment (Gavriilidou et al. 2019b).

Chapter 3

Large-scale cycling experiment

In the introduction chapter, the lack and need of empirical data in the form of cyclist trajectories were identified. In chapter 2, the approach to derive behavioural insights and to create models describing cyclist behaviour has been demonstrated with an existing dataset consisting of cyclist trajectories upstream of a traffic light. In order to extend the range of traffic situations and to overcome the data shortage, we performed a unique controlled, large-scale cycling experiment in the Netherlands. In this chapter the methodology for setting up and implementing such an experiment is described.These steps may be used as a guide in future experimental data collections and as a reference for future analyses using the data. Moreover, the collected dataset is introduced with an elaboration on its potential uses. The contribution of this chapter is, therefore, threefold: (i) it delineates the process to set up a large-scale cycling experiment; (ii) it describes the performance of the experiment, and; (iii) it presents the resulting large database of cyclist trajectories. Two subsets of the collected trajectories are used in the analyses and model development of chapter 4 and 5, respectively.

This chapter is based on the published article: Gavriilidou, A., M. J. Wierbos, W. Daamen, Y. Yuan, V. L. Knoop, S. P. Hoogendoorn (2019b) Large-scale bicycle flow experiment: setup and implementation, *Transportation Research Record: Journal of the Transportation Research Board*, 2673(5), pp. 709–719.

3.1 Introduction

Cycling as a main mode of transportation has in recent years been promoted by many governments worldwide due to its health and environmental benefits. The focus is mostly on finding ways to attract more people to the bicycle, while at the same time it is important to ensure a safe and comfortable infrastructure that can accommodate high cyclist volumes. This requires understanding of bicycle traffic characteristics, as well as insights into behaviour of cyclists while cycling on the road and making decisions to interact with other traffic participants and with the infrastructure. Research in this field is, however, limited and that is to a large extent due to the lack of empirical data.

To overcome this shortage of data, we performed a controlled large-scale cycling experiment. This chapter describes the methodology for setting up and implementing such an experiment. These steps may be used as a guide in future experimental data collections and as a reference for future analyses using the data. We describe the collected dataset and elaborate on its potential uses. The contribution of this chapter is, therefore, threefold: (i) delineating the process to set up a large-scale cycling experiment; (ii) describing the performance of the experiment, and; (iii) presenting a large database of cyclist trajectories.

This remainder of this chapter is structured as follows. In section 3.2 we provide a background of existing literature on operational cycling behaviour and identify the research gaps. Based on these, we formulate our research objectives in section 3.3 and discuss the findings of a stated preference survey that we conducted as a first step to meet the objectives (section 3.4). Section 3.5 describes the development of the data collection plan, while section 3.6 discusses its implementation. In section 3.7 the dataset is presented, followed by an outlook of future research.

3.2 Background on operational cycling behaviour

This section provides an overview of existing research on the operational cycling behaviour on an individual and on an aggregated level and identifies research gaps in each level.

3.2.1 Individual cycling behaviour

Operational cycling behaviour on an individual level can be represented by decisions regarding the use of the provided infrastructure while cycling and the interaction with other traffic participants.

In unconstrained situations, interaction decisions depend on the individual's choice for speed and positioning on the cycle path. A number of studies have looked into desired speed and acceleration profiles in free-flow conditions (Ma & Luo 2016; Twaddle & Grigoropoulos 2016), on different road surface types and gradients (Shepherd 1994), with normal bicycles as opposed to electric ones (Schleinitz 2016) and at wide or narrow cycle lanes (Vansteenkiste et al. 2013). These personal preferences might be constrained at high bicycle traffic volumes and when multiple directions intersect, an effect which is yet to be investigated. The interaction decisions in such situations, their coverage in literature and the corresponding knowledge gaps are the following:

- **steering to avoid colliding with other cyclists**: Steering manoeuvrers of bi-directional cyclists on collision course have been studied by Yuan et al. (2018), but the interaction with other directions is yet unknown.

- **overtaking cyclists**: Research on cyclists moving in the same direction has looked into following behaviour (Andresen et al. 2013), but overtaking decisions have not yet been investigated.

- **yielding to other cyclists**: To the best of our knowledge, there has been no research on yielding decisions at unsignalised crossings where priority rules apply, but are not enforced.

- **accepting a gap in a conflicting stream**: The gap acceptance of cyclists against right-turning vehicular traffic has been studied (Jiang et al. 2013). This, however, might differ significantly when cyclists interact with other cyclists and may also be influenced by whether the intention is to cross or merge.

- **stopping at a red traffic light**: Researchers have analysed red light running of cyclists at specific intersections across the world and identified influencing attributes that explain this behaviour, such as gender, age, amount of conflicting motorised traffic, crossing distance and cycling with company. An overview of these studies can be found in (Richardson & Caulfield 2015).

- **positioning when joining a queue**: The formation of multiple channels in queues has been observed at one signalised intersection, stressing the need for a bigger sample (Kucharski et al. 2017). The queue formation process in other situations, like upstream of an open bridge or a reduction of the cycle path width, is not yet studied.

3.2.2 Aggregated cycling behaviour

The aggregated behaviour of traffic participants is typically captured by the so-called fundamental diagram, which is the relation between average speed, density and flow. Several studies have investigated this relationship for cyclist flows and identified characteristics that are similar to vehicular traffic and pedestrian flow (Chen et al. 2013; Zhang et al. 2013). Other studies focused on understanding bicycle traffic flow and collected empirical data through:

- **single-file controlled experiments**: They have been conducted outdoors on circular tracks (Navin 1994; Andresen et al. 2013; Mai et al. 2013; Jiang et al. 2016; Zhao & Zhang 2017). In this setting, bicycle flow in low and high density situations can be observed, resulting in empirical data covering the full density range of the fundamental diagram. This provided insights into the dynamics of bicycle flow and identified flow characteristics such as stop-and-go waves. However, overtaking was not allowed in these experiments, which is often observed in real-life situations.

- **observing cycling behaviour in daily traffic**: Studies have estimated capacity of bicycle paths and resulted in a wide range of values (Botma & Papendrecht 1991; Li et al. 2015; Greibe & Buch 2016; Jin et al. 2017). This might be explained by the differences in infrastructure or bicycle type composition. The influence of electric bicycles has been studied (Wang et al. 2015; Zhou et al. 2015; Jin et al. 2017), but could not be controlled due to the nature of the empirical data. By controlling the infiltration rate of electrical bicycles, its impact to the overall flow characteristics can be identified more clearly. Furthermore, most empirical data is collected in conditions with low cyclist volumes and lacks observations in the congested regime of the fundamental diagram.

In short, the literature so far provides limited insight into the bicycle flow dynamics for high demand situations when overtaking is allowed and the effect of different attributes, such as the infiltration rate of electric bicycles, on the shape of the fundamental diagram, has not yet been studied.

3.3 Research objectives

Based on the given literature overview, it can be concluded that the research effort to observe and understand cycling behaviour is limited. The most essential gap seems to be studying high cyclist volumes, as well as bicycle-to-bicycle interactions at designated cycling infrastructure. With respect to individual behaviour, overtaking and yielding have been studied the least. At an aggregated level, overtaking is also important, as it is expected that it can explain the flow differences in the congested regime. Its effect on the shape of the fundamental diagram has not yet been studied, nor has the penetration of electric bicycles.

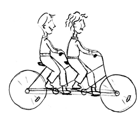
In order to address these gaps, we focus on bicycle traffic in the absence of other transport modes. Our objective is to collect a novel dataset that captures high cyclist volumes and where overtaking and yielding interactions take place. The aim of this dataset will, then, be to retrieve the characteristics of the fundamental diagram when overtaking is allowed and also to study the effect of bicycle type, an in particular electric bicycles, to the overall flow dynamics. Moreover, the dataset will be used to investigate the attributes that best explain the decisions to overtake and yield.

3.4 Survey on influencing attributes

To investigate the attributes that can explain overtaking and yielding decisions, we conducted an online stated preference survey in the Netherlands in summer 2017. The respondents were asked to name the attributes that influence their decision making in three situations: (i) overtaking or staying behind a single or a small group of cyclists; (ii) going ahead or stopping at a crossing to allow cyclists with priority to merge or cross, and (iii) stopping or continuing at a red traffic signal. The latter was included to check whether the attributes found from observations match the stated ones and as such justify the predictive value of the survey. The specificities of each situation were outlined, and always involved cycling during daytime on road infrastructure designated for cyclists and separated from other traffic. Per situation, a list of attributes was provided to the respondents based on behavioural hypotheses regarding the most influential attributes. Each list contained ten attributes displayed in random order, and three empty fields to enter other attributes. A selection of three to ten attributes was requested per situation. Apart from that, general information about the respondents was collected, such as gender and nationality.

By analysing the 444 responses, using principal component analysis to reduce dimensionality, the most influential attributes per decision could be obtained. In Figure 3.1 the prevalent attributes for each decision are linked to the corresponding decision (the check marks indicate the attributes that can be studied with our dataset). These decisions are part of the individual behaviour, together with steering and pedalling decisions. The schematic fits into the conceptual model of Figure 3.2 which describes cycling behaviour at the operational level. According to it, attributes influence the behaviour of individuals, who collectively give rise to aggregated behaviours. These behaviours can be observed via microscopic and macroscopic variables, whose relations are visualised in the conceptual model.

The validity of the survey findings is demonstrated by the attributes found significant for the decision to stop at a traffic light as they match those found in literature. However, more data are needed to quantify the effect of the attributes on the overtaking and yielding decisions. A data collection plan is, thus, necessary.



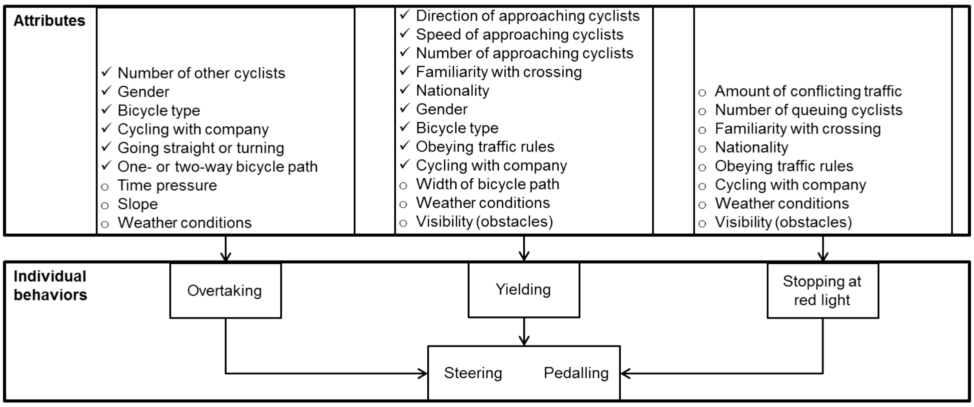


Figure 3.1: List of influential attributes per decision according to our survey results. The check marks indicate the attributes that can be studied with the dataset collected in our experiment.

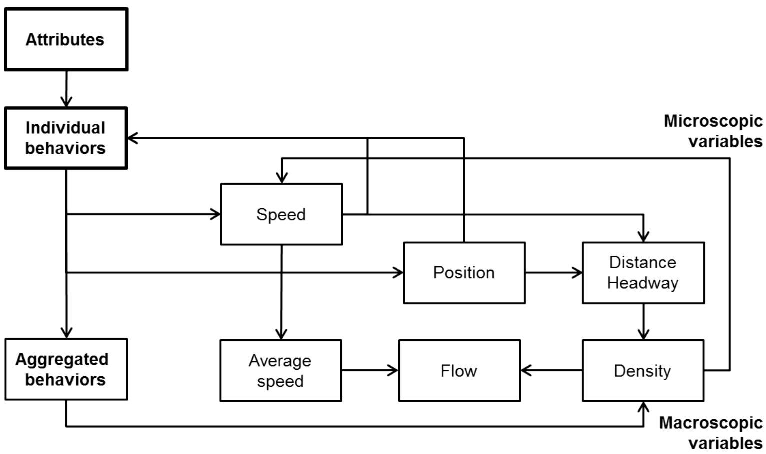


Figure 3.2: Conceptual model of operational cycling behaviour. Attributes are linked to individual behaviours, as already shown in detail in Figure 3.1. Collectively they lead to aggregated behaviours. These behaviours can be observed via micro- and macroscopic variables.

3.5 Development of data collection plan

The research steps to set up the data collection plan are described. First, the data needs and requirements are identified (subsection 3.5.1), followed by the motivation of the choice for the data collection approach and equipment (subsection 3.5.2). A controlled experiment is selected and its set-up is presented, covering the design of the scenarios (subsection 3.5.3) and the cycling track (subsection 3.5.4), the estimation of participants needed (subsection 3.5.5) and the duration required for each scenario (subsection 3.5.6).

3.5.1 Data needs and requirements

As previously mentioned, one of our aims is to retrieve the characteristics of the fundamental diagram when overtaking is allowed and the fleet consists of different bicycle type compositions, as well as to investigate the overtaking and yielding decisions of individuals and to identify the attributes that best explain them. The data type necessary to study individual cycling behaviour is trajectories, i.e., cyclist positions in time and a two-dimensional space. Trajectories are the most detailed type of traffic data, which can be aggregated in time or space to study macroscopic variables needed for the construction of the fundamental diagram. By examining trajectories, the use of the cycle path width and the speed adjustments can be studied relative to the position and speed of other cyclists and the environment (width, curve). The accuracy that is required for the trajectories lies within 10 cm which sets requirements for the data collection equipment. Additionally, it is crucial to be able to track and distinguish each individual, while also linking the observations to personal characteristics.

Another requirement is set by the need to capture the fundamental diagram. Therefore, it is necessary to observe low as well as high densities, which can be achieved by controlling the infrastructure setting and bicycle inflow rates. Studying the effect of different bicycle types means that the composition of the fleet should also be controlled. Moreover, controllability is necessary to ensure that the desired cyclist interactions (overtaking and priority negotiation) take place and that the effect of the influencing attributes of Figure 3.1 can be investigated.

3.5.2 Data collection approach and equipment

Three data collection approaches can be used to retrieve trajectory data:

- **observing real-life situations**: Even though this approach can capture the uninfluenced and unbiased behaviour, the degree of controllability is very low and does not meet the prescribed requirements.

- **doing an experiment in virtual reality**: Existing bicycle simulators are of unknown validity and behavioural realism. They also do not allow for multiple individuals to cycle simultaneously and interact with each other.

- **doing a controlled experiment in a physical environment**: A controlled experiment allows for a high degree of controllability and, thus, satisfies the requirements.

In a controlled experiment, the number of cyclists using the infrastructure, the routes they take, as well as the design of the infrastructure itself can be controlled. By carefully instructing the participants, specific elements of their behaviour, like their choice of speed, can be steered when necessary. Even the external conditions, such as light and wind, may be controlled.

However, the approach has some disadvantages that should be mitigated as far as possible through the experimental design. One of the main disadvantages is the occurrence of the so-called "learning-effect". This means that participants change their behaviour over time as their familiarity with the experimental setting increases and they get tired. This can be minimised by varying the layout and tasks that the participants are asked to perform during the day and by shortening their cycling duration.

Another potential drawback relates to data validity and representativeness. It may be argued that the behaviour is not realistic due to the fact that participants know they are being observed. We counter this argument based on the fact that the behaviour is observed several times and as they need to interact with other cyclists, their consciousness shifts to the riding task and any differences observed in behaviour are attributed to intra-personal variability. Moreover, this is intuitive behaviour, and the observation equipment will hardly be visible.

Regarding the data collection equipment, as the trajectories need to have high accuracy, overhead video cameras are selected. By placing them above the cyclists, the cameras track their movements with as little occlusion as possible, and continuously in time. In order to be able to automate the extraction of trajectories from the video images, a red cap is assigned to each participant. This is because red is the colour easiest to recognise under a wide range of lighting conditions (Daamen & Hoogendoorn 2003). Last but not least, since it is crucial to be able to link the observed trajectory to a specific individual, the caps are assigned a unique identification code.

3.5.3 Scenario design

On a microscopic level, the aim is to investigate the effect of the attributes of Figure 3.1 on overtaking and yielding decisions. In the scenario design we can control for two of them, namely the bicycle type and the directionality of the cycle path. Regarding bicycle type, separate runs are scheduled each with a different fleet composition and the scenarios are referred to as “Overtaking”. More specifically, there is a run for regular bicycles only, runs that combine regular bicycles with one special type, and a run with all types. In these scenarios there is a one-way flow on the cycle path. For the fleet with all types, the behaviour is compared with a run that allows for bi-directional flow.

With respect to yielding decisions, the direction of approaching cyclists is an attribute. Its effect can be investigated by separately studying crossing and merging streams. Therefore, two scenarios are designed, namely “Crossing” and “Merging”. As the bicycle type is an attribute, runs are performed with a mixed cycling fleet as well as with regular bicycles only.

On a macroscopic level, scenarios are needed to observe low as well as high densities to construct the fundamental diagram. We implement this by narrowing the cycle path, which obstructs the cyclist flow and leads to queue formation upstream of the narrow section when the demand exceeds its capacity. By varying the width of the narrow path (“bottleneck”), various congested patterns occur, determining both density and speed upstream of the bottleneck. We call it “Active bottleneck” scenario. It consists of different runs, each having another bottleneck width or a different cycling fleet composition, to observe the effect of bicycle types on the fundamental diagram. Specifically, the effect of electric bicycles is investigated by comparing three penetration rates: 0%, 10%, and 20%. These values represent typical values of electric bicycles in urban traffic situations in the Netherlands.

3.5.4 Track design

The layout of the track needs to be carefully designed because it largely determines the behaviour that can be observed in the experiment. First of all, cyclists should maintain a speed as close as possible to their normal cycling speed and behave as they would in reality. For this reason, a continuous track is selected, where participants make laps instead of short stretches that would require frequent acceleration from, and deceleration towards, standstill. A rounded rectangle shape is preferred over a circular one, because: (i) the cyclists will not be constantly steering in a curve; (ii) there is a straight stretch for overtaking manoeuvrers, and (iii) there is the possibility to study the effect of the attribute “going straight or turning” for overtaking decisions.

In terms of dimensions, the length of the straight stretch is set at 40 m, which is an adequate length for cyclists to overtake (Yuan et al. 2018). The width of the track is chosen to be 2 m. This width ensures that there is enough space for cyclists to overtake and it is also possible to sketch situations with a bi-directional flow (Zeegers 2004). The radius of the curve should allow cyclists to maintain a comfortable speed without the inside pedal hitting the surface if they lean. For a riding speed of 20 km/h, the minimum radius is 7 m (Shepherd 1994).

In order to ensure that the desired interactions take place, different track elements have been integrated into a single track layout, see Figure 3.3. The blue continuous line is the main track, used in all scenarios, where cyclists enter at the top left corner and cycle clockwise. The choice for this cycling direction is based on the norm to cycle on the right-hand side in the Netherlands and as such the inside curve will be taken by the slower cyclists. The inside curve radius is set at 10 m, with a quarter of a circle placed on each side and connected with a long straight stretch of 40 m and a short one of 16 m. The long stretch on the top side gives room for overtaking, while the bottleneck is placed at the bottom side in the Active bottleneck scenario. The short stretch accommodates crossing at a straight stretch rather than within a curve.

Another element is activated to observe crossing behaviour (green dotted line in Figure 3.3) where cyclists are riding counter clockwise. With this configuration, there is a bi-directional flow on the top part where the two routes overlap enabling the investigation of the effect of "one- or two-way cycle path" on the overtaking behaviour, and also creating two crossing points which increases the amount of observations. An extension of 10 m of straight stretch is added at the crossing points and the curve radius is set at 8 m, such that the crossing takes place in the middle of the blue track.

A third element is added (black dashed line in Figure 3.3) for the Merging scenario, which is connected to the main track in two locations; one is the off-ramp where cyclists can exit the main track and the other one is the merging point where cyclists join the main flow again. It is worth noting that no markings indicating priority are added on the track to prevent that they influence the behaviour.

With respect to controlling the flow, a bottleneck is introduced at the bottom side of the track. It consists of two inflatable mattresses placed next to each other on the track to create a narrow stretch 4 m long. The height of the bottleneck is 33 cm which blocks pedalling over it but does not hinder steering, creating the impression of an elevated curb rather than that of a wall which could be unsafe to drive through. The bottleneck is moved inwards to decrease the width of the track in that section. This way the cyclists are obstructed, leading to queue formation when the cyclist demand exceeds the capacity of the bottleneck. It is placed downstream the straight stretch (seen from the cycling direction) ensuring that the queue will grow along the straight stretch, and the observations are uninfluenced by the curve. By varying the bottleneck width, various congested patterns occur upstream of the bottleneck.

The bottleneck is set to four different widths, namely 75, 100, 125 and 150 cm. These numbers are based on a preliminary bicycle flow experiment that we performed, where the main path width was also 2 m and the path was narrowed to a width of 150 to 50 cm using steps of 25 cm. The 50 cm width was found to be too narrow for safety reasons. In order to observe high densities, the flow through the bottleneck should then be reduced in a different manner. The shape of the bottleneck is changed from a small straight stretch to one that cyclists have to meander through, referred to as the "Meander". The two mattresses are placed behind each other with 2 m space in between and in such a way that they leave a path of 75 cm to the side of the track (Figure 3.4).

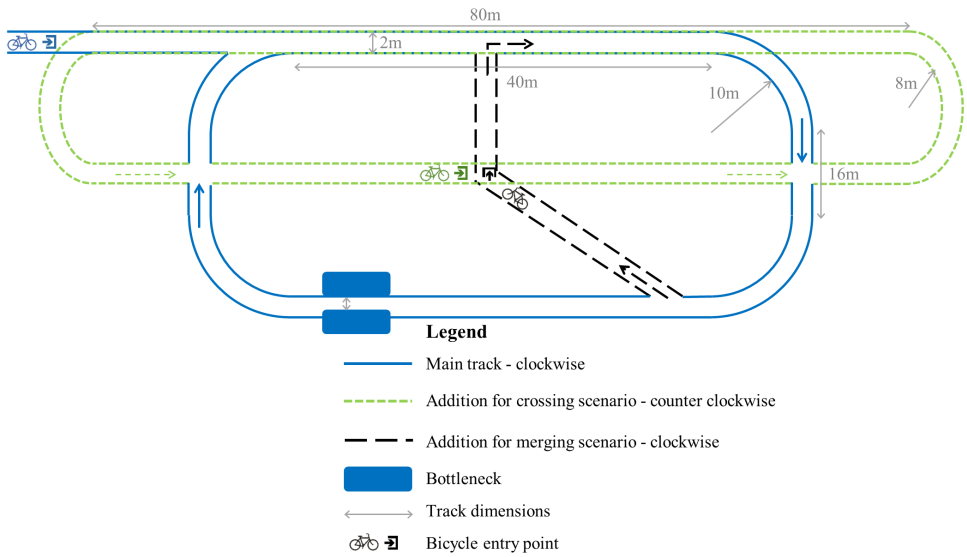


Figure 3.3: Track layout showing in colour the elements activated for different scenarios.

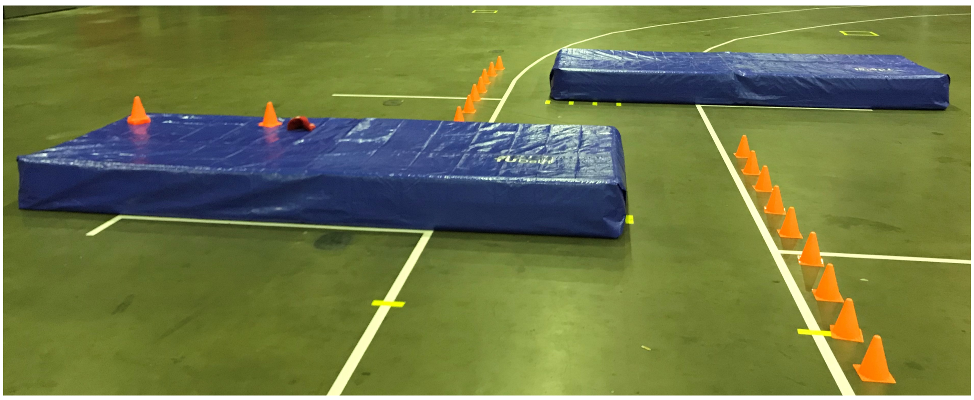


Figure 3.4: Construction of meander bottleneck using two mattresses.

3.5.5 Number of participants

The next step is determining the number of participants. We base this primarily on the aim to capture the relation between density, speed and flow. Assuming a diamond queue formation of 2-1-2-1, the jam density is 0.7 cyclists/m^2, which leads to 28 cyclists for a queue length of 20 m, which is enough to observe the behaviour for the high density and low speed situation.

To maintain a 20 m long queue, there need to be as many cyclists joining the tail of the queue as leaving the queue through the bottleneck. The number of cyclists that need to circulate the track depends on the outflow rate of the bottleneck, as well as on the average cycling speed of the circulating cyclists. To estimate the maximum number of participants, the scenario with the highest queue outflow rate should be considered. Based on our preliminary experiment, the outflow rate of the bottleneck of 150 cm width is 1.82 cyclists/s. Based on an average cycling speed of 19 km/h (Botma & Papendrecht 1991), 55 additional cyclists are needed. Consequently, a total number of 83 participants is required in the experiment.

3.5.6 Scenario duration and scheduling

The estimation of the duration needed for each scenario is based on the requirement to have enough observations to draw statistically significant findings. In the Overtaking scenario this is translated into giving each cyclist the chance to make at least ten decisions whether to overtake or not (i.e., cycle through the top straight stretch). When the bottleneck is inactive, it takes about 30 s to complete a lap, which leads to a required duration of 5 min.

With respect to the Merging and Crossing scenarios, the indicator to base the observation calculations on is the attribute “number of approaching cyclists”. In order to investigate its effect on the decisions being made, different group sizes, i.e. number of cyclists approaching the negotiation point from each side, need to be observed. As large numbers are appreciated the bottleneck that would constrain the outflow is removed. The time needed to collect sufficient observations of different group sizes is calculated using a simple microsimulation. It assumes a constant cycling speed and simulates dots moving around the track. Once a dot is detected close to the negotiation point, the number of dots present on each approaching stream is counted, while taking into account a physical length of about 2 m. If both approaches have a positive number, it is counted as an interaction of a group size coming from the right against a group size coming from the left. After running for a longer duration, the number of encounters of the occurring group combinations at the merging and crossing points is calculated. The result is that the Merging scenario requires 40 min and achieves interactions with a maximum group size of 6 against 5 cyclists, and every combination in between. Since the Crossing scenario has two observation points on the track, it requires half the time (20 min) for these observations.

In the Active bottleneck scenario, a 5 min duration is chosen. This duration enables the estimation of flow, density and speed in continuous and homogeneous conditions in the queue, without lengthening the total duration of the experiment. Also, it accommodates capacity estimation using different aggregation times, which decreases the influence of individual behaviour. Since participants are able to pass the bottleneck multiple times, approximately 5-10 times depending on the bottleneck width, the individual behaviour averages out which benefits the capacity estimation.

In terms of scheduling, the day of the experiment is divided into two sessions, one with special bicycle types and one without, so that we can observe the behaviour of regular bicycles only and compare it to the behaviour when special bicycle types are present. In the latter, the runs with these special types are dominant, checking the overtaking behaviour and the fundamental diagram for different penetration rates. Only two bottleneck widths (75 and 125 cm) are kept to limit the total running time. In the session without special bicycles, there is time to test all the widths and to focus on the Merging and Crossing scenarios. Due to the fact that the latter require long observation times, we split the duration in batches of smaller runs of 10 min each.

It is estimated that it takes 2 min for all cyclists to enter the track in a one-by-one pattern, and therefore the Overtaking and Active bottleneck scenario runs are scheduled to last 7 min. Since three fleet compositions (no special types, electric and regular bicycles, all types) are in both scenarios, their corresponding runs are scheduled in continuation, i.e., without any break. First the Overtaking scenario takes place and then the bottleneck is activated, which is estimated to take 1 min.

The activation is performed by introducing a moving bottleneck on the track, i.e., two persons cycling slowly and next to each other such that they cannot be overtaken and forming a queue behind them. This way, all cyclists are led as one group up to the bottleneck, activating it.

Summing up all these times leads to a net cycling time of 90 min for each session. To prevent exhausting the participants, breaks of 15 min are scheduled every 3 runs and in between runs there is a small pause of 5 min to initialise the next one. Apart from exhaustion, the learning effect and boredom need to be prevented. We solve this by alternating the scenarios in the schedule and by keeping the runs at about 10 min each. The planned order of scenario runs and their properties are summarised in Table 3.1.

Table 3.1: Schedule of scenario runs during the day of the experiment.

Scenario	Bicycle fleet composition				Bottleneck width	Run duration	Time to next run	Session
	Regular	Electric	Racing	Cargo	[cm]	[min]	[min]	
Overtaking	60%	20%	10%	10%	125	7	1	Morning
Active bottleneck	60%	20%	10%	10%	125	5	5	Morning
Overtaking	86%	-	14%	-	-	7	5	Morning
Active bottleneck	75%	25%	-	-	75	7	15	Morning
Overtaking	75%	25%	-	-	125	7	1	Morning
Active bottleneck	75%	25%	-	-	125	5	5	Morning
Merging	60%	20%	10%	10%	-	10	5	Morning
Active bottleneck	86%	14%	-	-	75	7	15	Morning
Crossing	60%	20%	10%	10%	-	10	5	Morning
Active bottleneck	60%	20%	10%	10%	75	7	5	Morning
Active bottleneck	86%	14%	-	-	125	7	15	Morning
Merging	60%	20%	10%	10%	-	10	5	Morning
Overtaking	86%	-	-	14%	-	7	-	Morning
Overtaking	100%	-	-	-	125	7	1	Afternoon
Active bottleneck	100%	-	-	-	125	5	5	Afternoon
Active bottleneck	100%	-	-	-	100	7	5	Afternoon
Merging	100%	-	-	-	-	10	15	Afternoon
Active bottleneck	100%	-	-	-	75	7	5	Afternoon
Crossing	100%	-	-	-	-	10	5	Afternoon
Active bottleneck	100%	-	-	-	150	7	15	Afternoon
Active bottleneck	100%	-	-	-	Meander	7	5	Afternoon
Merging	100%	-	-	-	-	10	5	Afternoon
Active bottleneck	100%	-	-	-	Meander	7	15	Afternoon
Crossing	100%	-	-	-	-	10	5	Afternoon
Merging	100%	-	-	-	-	10	-	Afternoon

3.6 Implementation of experimental design

Having set the requirements and the experiment design, the implementation follows and is divided into the selection of the location, the recruitment of participants and the set-up of the measuring and tracking equipment.

3.6.1 Location selection

The selection of the place where the experiment can be executed is based on several criteria. The most important criterion is that it has enough space to fit the track. The floor area required for the designed track is 100 m x 40 m. Moreover, the location should strictly prevent the presence of other modes. These conditions, along with the fact that a specific track with this shape and curves will be hard to find, point towards the construction of the track at a location rather than the use of existing infrastructure. Another benefit of creating the track is that it can be made obstacle-free to ensure good visibility. Even though the visibility due to obstacles has been found to be an attribute in the yielding decision, it is left out of scope to avoid accidents during the experiment.

Another criterion relates to the controllability of external conditions such as weather and light. These can only be controlled when the experiment takes place indoors. The weather conditions influence cycling behaviour, but investigating their effect would require repeating the experiment under different circumstances which is hard to predict and anticipate, as well as costly and difficult to plan with a sufficient number of participants. Therefore, we need to keep the circumstances constant during the whole experiment.

The indoor environment raises two needs. Firstly, the ceiling to be at least 10 m high to accommodate tracking equipment and prevent the feeling of cycling in a closed space. Secondly, the surface type should resemble real-world cycling conditions, be safe, and, therefore, be neither slippery nor adhesive.

Last but not least, the location should be easy to find and access, preferably near a crowded and inhabited area. This increases the chances of recruiting enough participants who will show up on time.

Given these criteria, we selected a large exhibition hall in the Ahoy Convention Centre, Rotterdam (The Netherlands). The size of the rented hall is 142 m x 70 m x 12 m, which satisfies all the dimension requirements and the floor surface is cement, so similar to cycling on road surface. Furthermore, it is well accessible by bicycle, connected by public transport and has a car parking.

3.6.2 Participant recruitment

The next step in the implementation is the recruitment of participants. Since it is desired to study the effect of gender and nationality of cycling behaviour, anyone is welcome to join. The only restriction is set with respect to age due to ethical reasons, and it is being at least 16 years old. A maximum age threshold is not set, but participants are asked to be physically able to cycle for around 90 min including breaks. As reward for the time they spent in the experiment, participants are given a small monetary compensation.

In order to increase the behavioural realism, participants are asked to bring their own bicycle. Upon registration, participants are asked for the bicycle type they intend to bring, as well as for other bicycle types they own. Special focus is placed in the recruitment phase on three special bicycle types (racing, electric and cargo).

Registration is performed through an online form, where availability in time of day (morning/afternoon session) and bicycles is declared. For those that meet the requirements, a confirmation is sent which includes the request to avoid red clothing which obstructs the tracking of the red caps in the camera images. Several platforms are used for the recruitment, such as posts in social media, universities and schools in Rotterdam and advertisements in local newspapers.

3.6.3 Measuring and tracking equipment

As previously mentioned, cameras are placed above the track to record the cyclist movements throughout the day. Due to the lighting conditions of the hall, which were low and variable, high quality cameras had to be used. Two snapshots of the experiment are shown in Figure 3.5. Figure 3.5(a) is a side view (from an overview camera, not to be used for tracking) during a Merging scenario. The cameras are placed at the ceiling next to the lights to improve the image quality and are 10 m above the ground. In order to cover the complete straight stretches, three cameras are required on each side with an overlapping area to ensure a continuous trajectory. Two more cameras are placed above the crossing points to observe the cyclist interactions there.

A top view at the location of the bottleneck can be seen in Figure 3.5(b). From this view the trajectories can be extracted by tracking the red cap of each cyclist. As shown in the image, each cap has a pattern of white boxes (like a bar code) on the flap which is unique and linked to the participant characteristics. An additional dot is marked in the middle, to identify looking and cycling direction.

Last but not least, we set up a corner to measure three main bicycle dimensions, i.e., full bicycle length, length from the front wheel to the handlebar and width of the handlebar. This enables studying the effect of different sizes on the behaviour in addition to the bicycle types.

(a) Side view of the Merging scenario.

(b) Top view of the Active bottleneck scenario.

Figure 3.5: Camera snapshots.

3.7 Experiment execution and high-level description of data

The experiment took place on 25 April 2018 with 178 participants evenly spread over the morning and afternoon sessions. This section presents the collected dataset, starting with adjustments of the plan that were needed during the day and continuing with the statistics of the participant characteristics and a qualitative description of the data.

3.7.1 Plan adjustments

During the first run in the morning session, it became clear that there were too many cyclists on the track. The queue configuration of 2-1-2-1 that was expected upstream the bottleneck was not observed. Instead, participants anticipated the bottleneck and started braking already at the curve. This resulted into a lower density than anticipated and an overall low speed (congested conditions).

The solution was to create two groups with half of the participants and alternate the group on the track. This way, the long breaks could be skipped as the participants could rest when the other group was cycling. Thanks to this change, it was possible to not only follow the plan, but have time for some additional scenario runs.

Since the narrowing at the bottleneck was anticipated and a dense queue was not naturally arising, we activated it using the moving bottleneck (i.e. the two persons in orange vests in Figure 3.6).

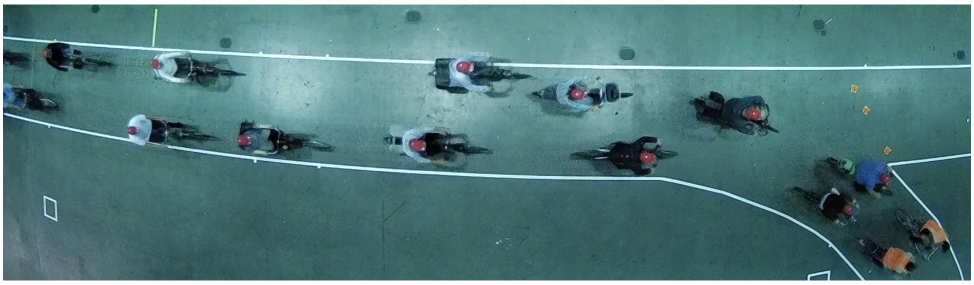


Figure 3.6: Queue formation behind a moving bottleneck.

In the Merging scenario, we initialised with a group starting from inside but the participants self-organised during the runs and dynamically shifted among the two routes. We decided not to obscure this process since it enhances observation of heterogeneity and could even lead to a model on route choice.

3.7.2 Participant characteristics

The descriptive statistics of the participants and their bicycles are summarised in Table 3.2 per session. It can be seen that more males participated in the experiment with a higher share in the afternoon session. The majority of the sample is Dutch and there is a wide range of ages.

With respect to the bicycles, the morning session contained special bicycle types with a high share of electric (35%) and a considerable share of 9% of racing bicycles. Unfortunately, no participants with cargo bicycles could be recruited. In the afternoon, almost all participants had regular bicycles. On average, the bicycle dimensions seem consistent between the two sessions.

Table 3.2: Descriptive statistics of participants and their bicycles per session.

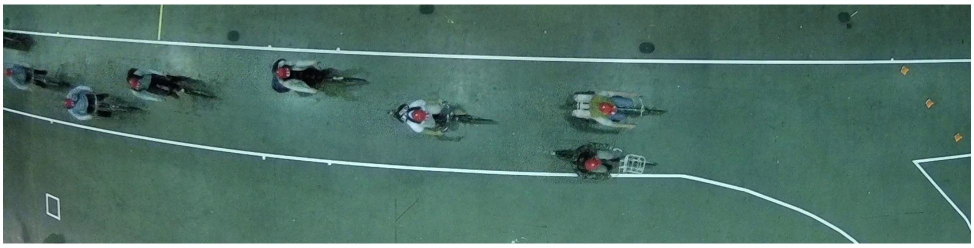
Characteristic	Morning session	Afternoon session
Females	34	30
Males	54	60
Dutch	78	84
Other European	8	2
Non European	2	4
Minimum age	19	17
Average age	52	51
Maximum age	80	89
Standard deviation of age	19	19
Average height [cm]	174	177
Standard deviation of height	10	10
Average weight [kg]	79	77
Standard deviation of weight	15	13
Electric bicycles	31	3
Racing bicycles	8	0
Average bicycle length [cm]	180	180
Standard deviation of bicycle length	6	5
Average handlebar width [cm]	59	59
Standard deviation of handlebar width	6	4

3.7.3 Qualitative data description

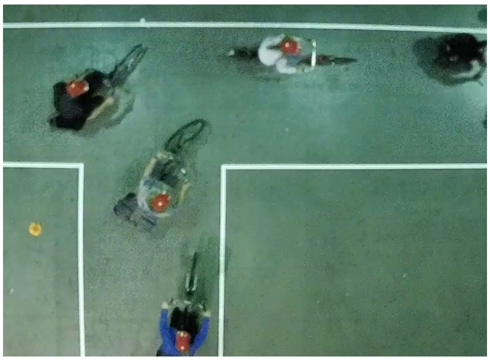
In total, six hours of videos have been collected which capture the cyclist movements throughout the day. Since there was time left, we tried one more situation. We had one run where we slowly filled up the track with everyone in to study the occurring wide moving jams. The planned scenarios were executed and additional to our expectations, the following phenomena were observed:

- Participants were braking already upstream the curve which led to lower than expected density.
- Many cyclists were overtaking in curves rather than the top straight stretch.
- Pairs were formed on the track, which blocked overtaking manoeuvrers (Figure 3.7(a)).
- A variety of yielding decisions was observed regardless of group sizes. Sometimes steering to create space was preferred to stopping (Figure 3.7(b)).
- During the Merging scenario runs, participants alternated between the two routes, leading to a dynamic share and different group sizes interacting at the merge.
- Some of the merging-route cyclists used their arms to indicate they would take the off-ramp and others taking that route would copy (Figure 3.7(c)).
- The right angle at the merging point was not always feasible to follow, so some cyclists went slightly off the track to merge (Figure 3.7(d)).

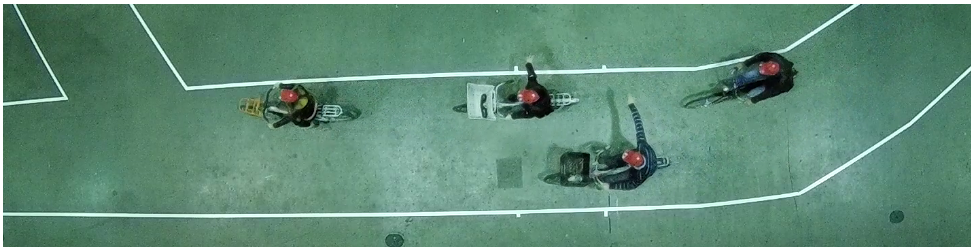
These observations show that anticipation plays a key role while cycling. In this obstacle-free environment where the curve and bottleneck were in sight, cyclists adjusted their speed in preparation for them. Moreover, speed differences could be better expressed in curves where the cautious cyclists would brake and the rest used this opportunity to overtake. Personal characteristics seem to be dominant with respect to yielding decisions and less so the number of approaching cyclists.



(a) Pair of cyclists obstructing the flow.



(b) Cyclist in black makes space for merging cyclists instead of yielding.



(c) Route indication using arms.

(d) Straying off the path to merge.

Figure 3.7: Examples of observed phenomena.

Participants respected the rule to cycle inside the cycle path, unless it would have led them to unsafe situations. Last but not least, self-organization has been found for cyclists in the form of distributing over different routes and copying the behaviour of others. These qualitative findings will be the starting point for future research, additionally to what was already intended with the collected dataset.

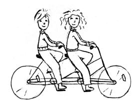
3.8 Future research with collected dataset

In this chapter, we described the set-up of a large scale controlled cycling experiment and qualitatively presented the collected dataset. The next research step is to process the video data and automatically extract trajectories out of the images, stitch trajectories between consecutive cameras and link to a participant number. This rich dataset will be used to investigate behaviour of different bicycle types and personal characteristics, and derive theoretical models that represent the decisions individual cyclists make while cycling and interacting with other cyclists, as well as models that describe the operationalisation of these decisions. The dataset will also be used to calibrate and validate these models. Apart from studying individual behaviour, we will study macroscopic bicycle traffic characteristics and construct the fundamental diagram for cycling. Such models can be used in future research to assess the quality of different bicycle infrastructure designs under several demand conditions.

Chapter 4

Cycling behaviour at T-junctions

In the previous chapter, the setup of a large-scale, controlled cycling experiment was described and the collected dataset was discussed qualitatively. In the current chapter, quantitative analyses and empirical findings from the experiment are presented. More specifically, the analyses focus on cycling behaviour at a bicycle T-junction. Intersections have been found to be critical locations in a network in terms of cyclist safety and T-junctions in particular have been found to lead to misperception of priority rules by car drivers. The analyses in this chapter investigate the extent to which this applies to cyclists. Additionally, the efficiency of bicycle flow at the T-junction is assessed. For this assessment a framework is developed that considers the effect of cyclist heterogeneity and of changes in the infrastructure design on the interactions that take place, on the use of the infrastructure and on the efficiency for each individual cyclist. In the collected dataset, several cyclist groups were observed, contributing to the heterogeneity analyses. Regarding the change in the infrastructure design, lane marking was added at the T-junction to guide the cyclist lane changing behaviour and facilitate the allocation of priority. The assessment framework of bicycle flow efficiency, the empirical findings related to cycling behaviour and efficiency, as well as the recommendations for the design of bicycle T-junctions are the main contributions of this chapter. The findings of this chapter are a stepping stone for the model developments in chapter 5 describing the yielding behaviour of cyclists.



This chapter is based on the published article: Gavriilidou, A., W. Daamen, Y. Yuan, N. van Nes, S. Hoogendoorn (2020b) Empirical findings on infrastructure efficiency at a bicycle T-junction, *Physica A: Statistical Mechanics and its Applications*, p. 125675.

4.1 Introduction

The increased interest and promotion of cycling by governments and municipalities throughout the world goes hand-in-hand with the provision of better cycling infrastructure. The design of this infrastructure is of great importance, because it influences cyclist behaviour, as well as the attractiveness to use it. Moreover, the design should offer a safe and efficient transition through the network. Critical locations in the network in terms of safety and efficiency are discontinuities of the infrastructure, obstacles placed within the infrastructure and locations where different transport modes or traffic streams interact and negotiate the use of space.

At intersections there is a higher risk for cyclists to be involved in a collision (Dozza & Werneke 2014; Flower & Parkin 2019). At signalised intersections, this is mostly attributed to rule violations by the cyclists or the motorised vehicles. Especially in Asian countries, the cause of bicycle crashes at signalised intersections has been found to be the red-light running and retrograde behaviours of electric bicycles (Pai & Jou 2014; Wu et al. 2012). However, the label of electric bicycles is also used for scooters (Fishman & Cherry 2016), which explains why red-light running behaviour of cyclists is not found so dangerous in Western world countries (Pai & Jou 2014). At unsignalised intersections, on the other hand, this higher risk may relate, among other things, to the presence of sharp turns and the lack of road marking or lighting (Wijlhuizen et al. 2016). Another cause might be the high number of interactions. A study in Amsterdam showed that when the number of cyclists at an intersection is high, the level of stress and discomfort of the cyclists rise, and they adhere less to traffic rules (Imbert & te Brömmelstroet 2014).

One type of intersections are T-junctions, which consist of three road segments (also known as arms) two of which belong to a straight road. In car traffic, it has been found that uncontrolled T-junctions can pose a problem when priority is assigned to the intersecting (right) arm, because drivers on the straight road have a high perception of priority and fail to yield (Helmers & Aberg 1978). Priority perception is very important for the type of interactions that occur. When priority perception is high, drivers as well as cyclists tend to have higher speeds and minimal head movements to observe their surroundings, which can result in unsafe interactions (Costa et al. 2019). These studies, however, investigated either only car traffic or mixed car-bicycle flows, where the priority was indicated by road signs. Given the findings of Wexler & El-Geneidy (2017) where it is suggested to keep the different transport modes in separated infrastructure, the transferability of the findings to dedicated bicycle T-junctions becomes questionable. Moreover, the effect of design possibilities other than road signs to indicate priority has not been researched.

We focus on T-junctions dedicated to cyclists and investigate cyclist behaviour using empirical trajectory data collected during a large-scale cycling experiment (Gavriilidou et al. 2019b). The reason for using observations from an experimental setting rather than real-world, is that there is control over the infrastructure design, the characteristics of the participants and their bicycles (Hoogendoorn & Daamen 2005). During the experiment, lane marking was implemented, aiming to guide through cyclists to the left and leave the right side of the path for merging cyclists. The hypothesis was that this lane marking could inflict separation of conflicting streams in space, leading to safer interactions and more efficient flow. The safety effects are presented by Nabavi Niaki et al. (2019). In the present chapter, we study the effect on infrastructure efficiency. In this context, the term efficiency is used to represent the extent to which the infrastructure can be used in relation to its design and the behaviour of its users. For the assessment of efficiency we develop a comprehensive framework, which is the first contribution of this chapter. Another contribution are the empirical findings we derive on the infrastructure efficiency of a bicycle T-junction. Based on these findings, our third contributions is that we make recommendations for the design of bicycle T-junctions.

The remainder of this chapter is structured as follows. In section 4.2 a comprehensive framework is described for the assessment of the T-junction efficiency and in section 4.3 the data from the controlled experiment are presented. The method to apply the framework is described in section 4.4. The results are provided and discussed in section 4.5, followed by conclusions and recommendations for the infrastructure design in section 4.6.

4.2 Infrastructure efficiency framework

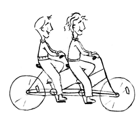
This study aims to assess the efficiency of bicycle infrastructure. Our focus is on T-junctions dedicated for bicycle use. In order to evaluate efficiency, a comprehensive framework is proposed that turns the most relevant influence factors into specific key performance indicators. This framework is depicted in Figure 4.1.

Infrastructure efficiency depends on the infrastructure design, as well as on the behaviour of its users. The design captures elements like the marking, the merging angle, the width and surface of the cycle path. Marking may be used to indicate which direction has priority. This is usually indicated by “shark’s teeth” (i.e. white isosceles triangles) drawn on the surface of the path, at the end of the minor (i.e. non-priority) stream or by road signs. When such markings and signs are missing, general traffic rules apply. Marking can also be used to exemplify the expected space utilisation, by drawing lines on the surface. Solid lines generally indicate a compulsory movement, while dashed lines are suggestive, for example to guide minor stream cyclists into the through lane. The merging angle may affect the perception of priority as well as the speed of the merging cyclists and the ease of their merging manoeuvre. The width of the cycle path affects the ability to form lanes as well as the overtaking behaviour. Last but not least, the type of surface may affect the cycling speeds and the perception of priority.

The term “cycling behaviour”, in this case, encompasses different behaviours, like the choice of speed and lane, distance keeping, lane changing, yielding and gap acceptance. This behaviour is not only influenced by the infrastructure design, but also by the traffic rules and the person’s attitude to adhere to them. Other personal (age, experience) and bicycle characteristics (type, condition) also have an effect.

The interaction between infrastructure design and cyclist behaviour determines the infrastructure efficiency. We hypothesise that efficiency cannot be captured by a single metric, but is rather assessed by a compilation of indicators. The proposed framework aims to comprehensively cover different aspects of efficiency. These aspects are grouped in three categories with increasing order of complexity. The first refers to single cyclists, the second to the use of the infrastructure, and the third to the interactions between cyclists.

Within the first group we consider the speed upstream and down-



stream of the T-junction, and the travel time. The use of the infrastructure in the second category refers to the capacity of the T-junction and the utilisation of the two-dimensional space. The capacity can be determined from the longitudinal time-headways that occur downstream the junction, while the space utilisation is reflected by the lanes that are formed in the lateral direction. The last group of indicators focuses on the cyclist encounters in conflicting directions. The group sizes that arrive upstream of the T-junction negotiating priority influence the delay that is inflicted. This delay can also be captured by the number and duration of stops that cyclists make to give way to the priority stream.

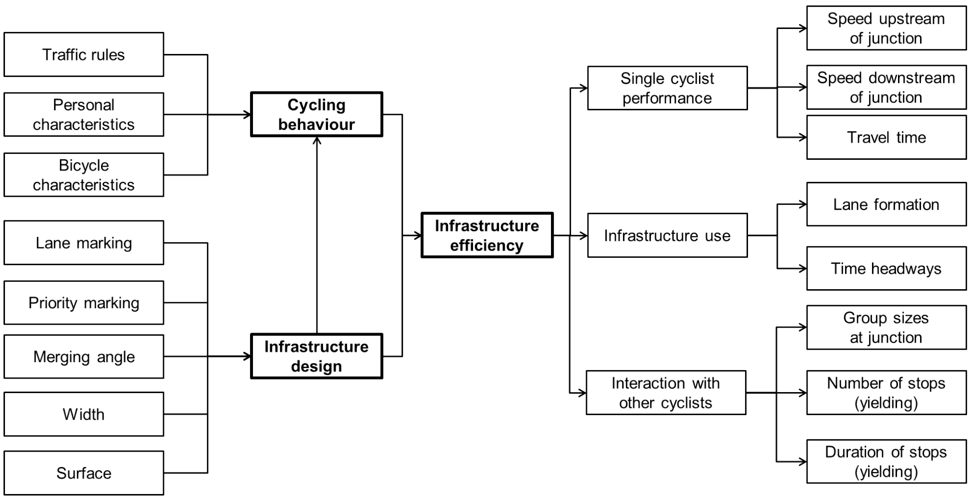


Figure 4.1: Relation between cycling behaviour, infrastructure design and infrastructure efficiency of T-junctions.

4.3 Cycling data at T-junction

The dataset used to assess the T-junction efficiency consists of cyclist trajectories collected during the controlled experiment presented in chapter 3. In this section, relevant details about the experiment are provided (subsection 4.3.1), followed by the description of the trajectory extraction (subsection 4.3.2) and a summary of the characteristics of the final dataset for which the key performance indicators are calculated (subsection 4.3.3).

4.3.1 Controlled experiment

This chapter focuses on one hour of video data from one of the cameras, namely the one placed over the T-junction. The selection of the one hour is such that it covers the scenario runs during which the route on the track that made use of the T-junction was open. These runs will be referred to as the "merging runs". During the whole experiment, there were six merging runs of about ten minutes each (hence the one hour of video), which were spread throughout the day to keep the schedule of the participants interesting, alternating between different routes and interactions. For details regarding the experiment, the reader is directed to Gavriilidou et al. (2019b).

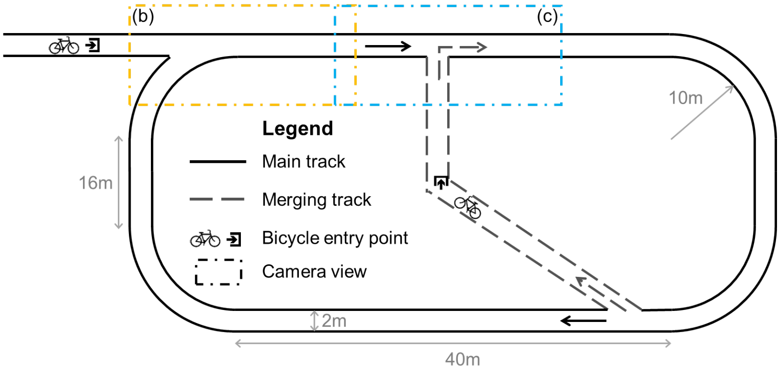
An overview of the track layout that was used during the merging runs is shown in Figure 4.2(a). In each run, 40-60% of the participants were asked to follow the main track route and the others followed the merging track route. The track was always marked with a continuous tape, so it is only for readability that the merging track is indicated in the figure with a dashed line. The camera view of interest for this study covers the T-junction and is indicated by the blue box in Figure 4.2(a). No other marking was placed on the track except for an infrastructural nudge that was implemented prior to the last merging run in order to study the effect of lane marking on the cycling behaviour at the T-junction. The lane marking consisted of a dashed line, starting at the edge of the track upstream of the T-junction and leading toward the centre of the track, continuing till downstream the T-junction as a centre line. A snapshot of the lane marking is shown in Figure 4.2(b) and Figure 4.2(c), which correspond to the camera views of the camera placed on top of the junction (blue box in Figure 4.2(a)) and the camera upstream of it (orange box in Figure 4.2(a)).

As shown in Figure 4.2(a), the two camera views overlap. This is to facilitate the coupling of trajectories extracted from different cameras. The boxes in Figure 4.2(b) and Figure 4.2(c) were marked with yellow tape on the ground, while the yellow letters A and B are added in the figure to demonstrate which boxes are the same in the two snapshots. The marked cross-sections in Figure 4.2(c) are not present on the ground, but are added to the figure to assist in the explanation of the calculation of the infrastructure efficiency indicators in section 4.4.

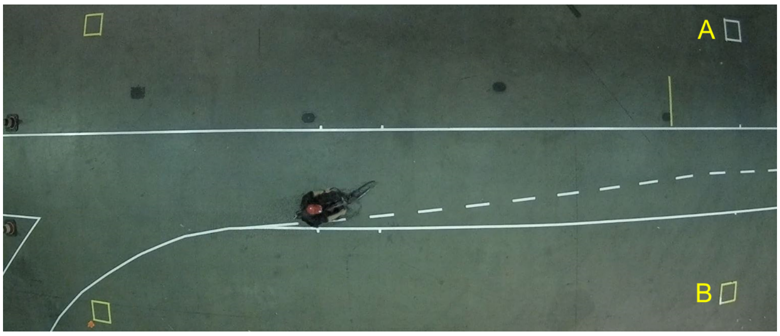
The aim of the lane marking is to direct through cyclists to the left lane and achieve separation of the conflicting flows at the junction. Ideally, this design would alleviate the need for through cyclists to stop, as they can freely continue into the left lane, while merging cyclists enter at the right lane. The through cyclist seen in Figure 4.2(b) indeed steers away from the dashed line.

Another observation concerns the design of the T-junction in Figure 4.2(c). The width of both streams was kept constant at all time at 2 m, which is enough for two cyclists to comfortably ride side-by-side and execute overtaking manoeuvres. The turning angle was also kept constant at 90 degrees. The motivation behind the choice of perpendicular cycle paths was to create a clear sense of priority, to be assigned to cyclists coming from the right. No priority marking was added to further influence the behaviour. Moreover, the taped edges of the track were chosen to be perpendicular rather than curved in order to observe to what extent cyclists cut-off the edge and identify the curve they find most comfortable for turning. This choice was made also for practical reasons, as a straight angle is easier to tape. The feasibility of the turn was tested and approved, so it was possible to execute the merging manoeuvre without straying off the track edges. Figure 4.2(c) shows a snapshot of a merging cyclist cutting the corner.

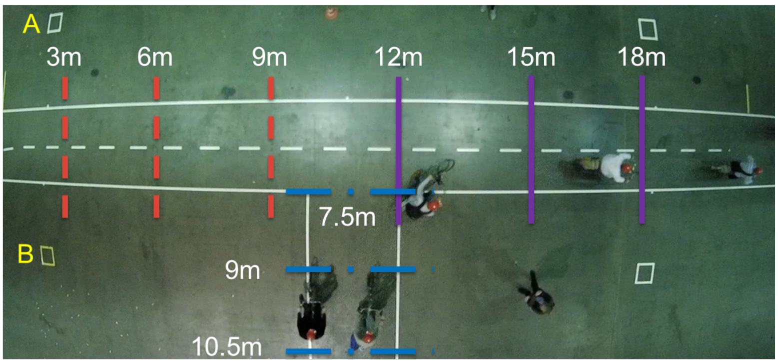
Regarding the cycling behaviour, participants were instructed to cycle as they would normally do, and to try to stay within the edges of the cycle path as much as possible. The given instruction left the choice of speed, priority allocation and overtaking to the participants. This means that they were neither forced to obey the traffic rules nor prohibited from overtaking other cyclists. The tape was thick enough to let passing cyclists know that they are crossing it, but not raised as a curb that could cause safety incidents.



(a) Layout of experimental track.



(b) Camera view upstream T-junction.



(c) Camera view at T-junction.

Figure 4.2: Experimental track layout and camera snapshots at the T-junction and upstream of it during the run with the lane marking.

4.3.2 Trajectory extraction

During the experiment, all participants wore a red cap which could be tracked in each frame of the videos, thereby allowing us to reconstruct the trajectory of each cyclist. The extraction of the trajectories followed six automated steps:

1. Identification of red caps in each video frame based on detection of red coloured pixels and clustering them into objects as described by Hoogendoorn et al. (2003).

2. Connection of centre points of the identified red caps in the consecutive video frames to construct cyclist trajectories using an adjusted version of the MODT-2 software developed by Duives (2016).

3. Transformation of head trajectories to ground trajectories by correcting for the height difference between the cap and the ground. This process made use of height at which the camera was placed (i.e. 12 m), and of the average height of persons on a bicycle which was measured to be 179 cm.

4. Orthorectification of the ground trajectories to correct for the camera distortion and convert from pixel units to metres. The correction is based on parameters determined by ImageTracker (Knoppers et al. 2012).

5. Removal of short, misidentified trajectories by ensuring that each cyclist is observed at two cross-sections (3 m and 18 m lines in Figure 4.2(c) for through cyclists and 10.5 m and 18 m for merging cyclists).

6. Replacement of blobs (i.e. clouds of points) observed during a stop by a single stopping location, thereby ignoring the movement of the head when the bicycle stands still.

4.3.3 Run characteristics

Table 4.1 shows an overview of basic characteristics of the merging runs. The first three runs took place in the morning session of the experiment, where participants could bring regular bicycles as well as racing and e-bikes. The 88 participants who joined this session were divided into groups, referred to as “Mix 1” and “Mix 2” because of the mixture of bicycle types. In the afternoon sessions, the 90 participants, who were asked to bring regular bicycles only, were split into the groups “Pure 1” and “Pure 2”. The last run in the afternoon is the run with the lane marking.

Since the identified trajectories do not correspond to specific person IDs, the number of trajectories shown in the 5th column of Table 4.1 does not reflect the number of participants, but rather the number of passages through the camera view. The share of identified merging trajectories ranges from 45% to 56%, which does not coincide with the share of cyclists taking the merging route, because the shorter distance of this route and the cycling speed affect the number of laps they make and, hence, the number of cyclists seen by the camera per run. The share of merging cyclists is not known, as it varied during the run, because participants took the liberty of alternating between the two routes (i.e. the main and merging tracks). These identified shares of merging trajectories can be seen as an outcome of the efficiency of the T-junction, rather than a cause. For this reason they are not further discussed or used to assess the efficiency. Instead, the assessment of the infrastructure efficiency is based on the framework presented in section 4.2 and corresponding indicators (to be introduced in section 4.4).

It can be seen that the number of trajectories differs per run, but so does the duration of the run. Therefore, the comparison should be based on the ratio of the two, which is the average flow of cyclists and is shown in the last column of the table. The flow values in all runs are around 1 cyclist per second and are, thus, comparable. Therefore, it is considered that the different number of trajectories has no effect on the findings of this chapter.

Having six runs allows investigating not only the effect of the lane marking, but also of the different bicycle types (“Mix” versus “Pure”) and persons (morning and afternoon, and group 1 versus group 2). Moreover, the results of the 3rd run compared to those of the 1st reveal whether any learning effect takes place as participants get familiar to the experimental conditions. The conclusion on whether there was learning effect is then used in the comparison between the 6th and 4th run, to separate the impact of the lane marking, from that of the learning effect. The impact of different bicycle types is found by comparing runs 1 and 2 with runs 4 and 5, while the impact of different personal characteristics is investigated through the comparison of run 1 with run 2, and run 4 with run 5.

Table 4.1: Basic characteristics of runs.

Run	Session	Bicycles	Lane marking	Number of trajectories	% Merging	Duration of run [min]	Flow [cycl/s]
1	Morning	Mix 1	No	480	50.8	8.8	0.91
2	Morning	Mix 2	No	417	46.8	6.6	1.05
3	Morning	Mix 1	No	629	56.1	10.0	1.04
4	Afternoon	Pure 1	No	598	55.9	10.1	0.99
5	Afternoon	Pure 2	No	610	45.3	10.0	1.02
6	Afternoon	Pure 1	Yes	803	51.8	13.1	1.02

4.4 Method to calculate the efficiency indicators

In section 4.2 the framework to assess the efficiency of a T-junction was presented. In this section, the method to calculate each of the efficiency indicators is described. The extracted trajectories have been the basis for the calculation of all the metrics. The results of the calculations are discussed in section 4.5.

4.4.1 Single cyclist performance indicators

Speed distributions are constructed from the speeds of the individual cyclists in a specific segment. A speed measurement is made based on the passing time, t, between two cross-sections in Figure 4.2(c). For example, for the through cyclists upstream of the T-junction, the cross-sections at 3 m and 9 m are used and the time instances at which each cyclist passed them (t_3 and t_9, respectively) are known from the trajectory data. The ratio of the distance covered and the corresponding time interval define the average cycling speed, v, for that stretch:

$$\begin{aligned} v_{\text{upstream,through}} &= \frac{9-3}{t_9 - t_3} \\ v_{\text{upstream,merging}} &= \frac{10.5-7.5}{t_{7.5} - t_{10.5}} \\ v_{\text{downstream},i} &= \frac{18-12}{t_{18} - t_{12}}, \text{for } i = \text{through, merging} \end{aligned} \tag{4.1}$$

If the infrastructure is efficient, the cycling speeds should be similar for the up- and downstream stretches of the same cycling direction, such that the cyclists in each stream can pass the T-junction uninterrupted (i.e. keeping the same speed). However, as the streams are intersecting, cyclists should also ensure safe interactions which require anticipation and deceleration. The infrastructure design that requires the least amount of deceleration is, thus, the most efficient. The expected effect is further elaborated for each stream. Through cyclists are expected to yield to the merging cyclists who have priority. The yielding through cyclists who come to a complete stop (i.e. zero speed) are not taken into account, as their speeds are very different in the two stretches and the properties of their stops are considered in other indicators. The cyclists who decelerate are included in the analysis as they show the effect of the interactions with merging cyclists on their speed. This effect should ideally be minimal, thus similar speeds upstream and downstream. Regarding the merging cyclists, it is expected that they slow down to make the turn. Any differences observed between the two stretches are, then, depicting the efficiency deterioration of the infrastructure design (i.e. the 90 degree turning angle).

Travel time is also calculated as a measure of delays. Its variability across the runs can reveal any deviations from the free-flow travel time. Since the latter is not known, the average of the six runs is taken as a proxy. The travel time, tt, is defined as the time interval during which a cyclist is present between two cross-sections:

$$\begin{aligned} tt_{\text{merging}} &= t_{18} - t_{10.5} \\ tt_{\text{through}} &= t_{18} - t_3 \end{aligned} \tag{4.2}$$

4.4.2 Infrastructure use indicators

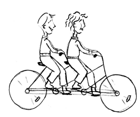
Lane formation and *time headways* are the two metrics for the use of the infrastructure. The former can be observed through a heatmap of the space utilisation, and the latter through the time interval between consecutive cyclists at a selected cross-section.

According to the standard dimensions of the width of a bicycle handlebar CROW (2016), space is discretised in squares of 67 cm. A maximum of three sublanes can then be formed and observed in the 2 m width of the cycle path. Using squares reduces the complexity of having to rotate the shape with the changing direction of the merging cyclists.The heatmap is constructed by counting the number of cyclists (trajectories) that ride over each square. A more efficient infrastructure would ensure that the entirety of the cycle path width is used.

The (minimum) time headways are a proxy for the capacity of the T-junction (capacity equals one over the average minimum time headway) and should, therefore, be measured downstream of it. The cross-section at 15 m (Figure 4.2(c)) is chosen and the time interval between consecutive cyclists passing it is recorded. The shorter the time headways, the higher the capacity and, thereby, the higher the efficiency of the T-junction.

4.4.3 Indicators for interaction with other cyclists

Under the assumption that the priority stream remains unaffected by the presence of through cyclists, it is the latter that need to react, i.e. either make a stop or reduce their speed or make space or accept a gap. The indicators reflecting the interactions with other cyclists have as starting point the presence of a through cyclist upstream of the T-junction. In order to consistently compare the interactions that each through cyclist faces, a cross-section needs to be selected as the critical moment for the decision of the through cyclist to react, if necessary, to the anticipated conflict. To the authors' knowledge there is no literature investigating how far upstream cyclists anticipate on conflicts and when they make their decisions. In this study, we take the cross-section at 6 m (Figure 4.2(c)) as the critical decision moment. One argument for this choice is that it is four metres upstream of the cross line for through cyclists (thus upstream of the intersection), which is the same



distance as that available in the camera view for merging cyclists. Another argument is that it is in the middle of the range for which the speed distributions have been investigated. The start of this range at 3 m is considered too far upstream for proper anticipation, while the 9 m at the end of the range are too close to the cross line which is at 10 m, and there is not sufficient reaction time if cyclists would make their decision there. The choice for the 6 m cross-section is further evaluated in the discussion of the results in section 4.5.

When a through cyclist passes the 6 m line, it is checked whether any merging cyclist is present upstream of the junction. If so, the number of cyclists present in each stream is counted and this represents an instance of two group sizes, one for the through stream and one for the merging. This process is repeated every time a through cyclist passes the 6m line, leading to a frequency of different *group size* combinations to occur. Given the length of the visible stretch for through cyclists upstream the T-junction, which is 10 m, the 2 m width of the cycle path and the average bicycle length of 2 m, a maximum of 10 through cyclists can be observed. Similarly, given the length of 4 m visible in the upstream merging stretch, a maximum of 4 merging cyclists is expected. In an efficient design the group sizes should be larger for the merging than for the through cyclists. The reasoning behind this is that if many through cyclists yield for only few merging cyclists, more delay is experienced than time savings.

As already stated, one of the possible reactions of through cyclists is to stop, yielding to merging cyclist(s). These stops are a form of delay and two metrics are defined to assess their effect on efficiency. The first one is the *stop duration*, corresponding to the time a through cyclist remains stationary. The second metric is the *stopping frequency.* This is a normalisation of the number of stops in the run to remove the effect of the run duration. The applied normalisation is the inverse of the time difference between two consecutive stops, so that the variation of stops within each run is captured as well. More specifically, the stopping frequency, S, in stops per minute is calculated for each stop, k, using the following equation:

$$S_k = \frac{60}{t_k - t_{k-1}} \tag{4.3}$$

The efficiency is higher when stops are shorter, which is linked to smaller group sizes and when there are less stops per minute. High numbers indicate that the interval between consecutive stops is high, so many cyclists need to yield at the same time or the outflow from the queue goes in waves and the same cyclist needs to make several stops.

4.5 Results and discussion of efficiency

The results for each set of indicators of the assessment framework in Figure 4.1 are presented in the following subsections, along with their discussion.

4.5.1 Efficiency based on single cyclist performance

Figure 4.3 shows the cumulative distribution functions of the average speed per run, direction and section of the infrastructure. The top row corresponds to the section upstream of the merging location, and the bottom to the downstream section. The left column is for through cyclists and the right for merging cyclists. The y-axis shows the cumulative probability and the x-axis the cycling speed through the corresponding segment in km/h. Runs with the same group of cyclists have been assigned the same colour and different line styles. Full lines correspond to the first run of that group of participants and dashed lines to the second run, when applicable.

It can be seen that the through cyclists in the upstream section have the lowest speed, which can be attributed to the fact that they sometimes need to decelerate and give way to merging cyclists coming from the right. Downstream of the junction, the through cyclists reach twice as high speeds. The opposite phenomenon occurs for the merging cyclists. Upstream of the junction they have higher speeds than downstream. This is due to the sharp right turn, when merging cyclists try to enter the right lane, which requires a speed reduction.

Kolmogorov-Smirnov tests are performed to compare the empirical distributions against each other and draw conclusions regarding the presence of a statistically significant difference (Dodge 2008).

Four comparisons are performed between the distributions of:

- **Runs of through cyclists:** When comparing the different runs to each other, it is found that there is no statistically significant difference at a 5% significance level between runs 1 and 3 which have the same population. This means that there is no evidence for a learning effect due to the experimental conditions, so through cyclists do not change their speed upstream of the T-junction as they get more familiar with the infrastructure and the occurrence of encounters. Contrary to this, the speed distributions in runs 4 and 6 are found to differ. Therefore, the lane marking appears to have an effect on the speed of the through cyclists upstream of the junction. The effect is positive, as more observations have higher speeds in the case of run 6. Another finding is that the upstream speed distributions of runs 5 and 6 are not statistically significantly different, which shows that the characteristics of the cyclist population (i.e. cyclist heterogeneity) are at least as important as the presence of lane marking. This heterogeneity is further attributed to personal characteristics, rather than the different bicycle types, as the speed distributions of runs 1 and 4 have no statistically significant difference. Similar results are found when comparing the speed distributions of the through cyclists downstream of the T-junction. The only difference is in the comparison of runs 5 and 6, which are statistically significant different downstream the T-junction. In this stretch, the cyclist heterogeneity leads to higher speeds than the presence of the lane marking.

- **Runs of merging cyclists:** Most speed distributions of merging cyclists are statistically significant different at a 5% significance level. Heterogeneity is observed in terms of personal characteristics as well as in the different bicycle types. Personal characteristics seem to be more prominent, as the presence of faster bicycles in the mix (i.e. racing and electric bicycles) does not necessarily result in higher speeds (run 4 with only regular bicycles is the fastest run, but run 2 with a mix of bicycle types is faster than run 5 with only regular bicycles). In this case, there seems to be a learning effect, as the "Mix 1" cyclists have higher speeds both upstream and downstream in the case of run 3 compared to run 1. This means that merging cyclists become more comfortable with the right turn and develop a strategy to take it with the least speed reduction possible. At the same time, and despite the positive contribution of the learning effect, merging cyclists are worse-off by the introduction of the lane marking, as they cycle slower in run 6 compared to run 4. The explanation might be the introduction of the lane separation downstream: they feel obliged to merge into the right lane without using the left lane that has been assigned to through traffic by the lane marking, so they are forced to decelerate already upstream. In other words, the lane marking makes the 90 degree turn even sharper.

- **Merging and through streams:** An interesting observation from Figure 4.3 is the complementary effect between the speeds of the two directions; the stream that is fastest in a run, is the slowest in the other direction of the same run. For example, run 5 was the fastest for through cyclists and the slowest for merging cyclists. This is more evident in the upstream section and has its roots in the way that cyclists handle their encounters. When merging cyclists are very fast, through cyclists experience very small gaps, which they cannot accept. So they have to yield until a sufficient gap becomes available, resulting in reduced average speed in the upstream segment.
- **Upstream and downstream stretches:** The comparison of speed distributions of the same run between the two stretches shows that they are all different at a 5% significance level. The speed in one stretch is almost half of that in the other for both cycling directions. This large difference in speed shows that the efficiency is low at the T-junction. Regarding the merging cyclists, it shows that the 90 degree is not an efficient infrastructure design. The same holds for the lane marking in the through direction, as it does not achieve that the speeds upstream and downstream the T-junction remain similar. The through cyclists have quite some interactions where they need to decelerate and give priority to merging cyclists.

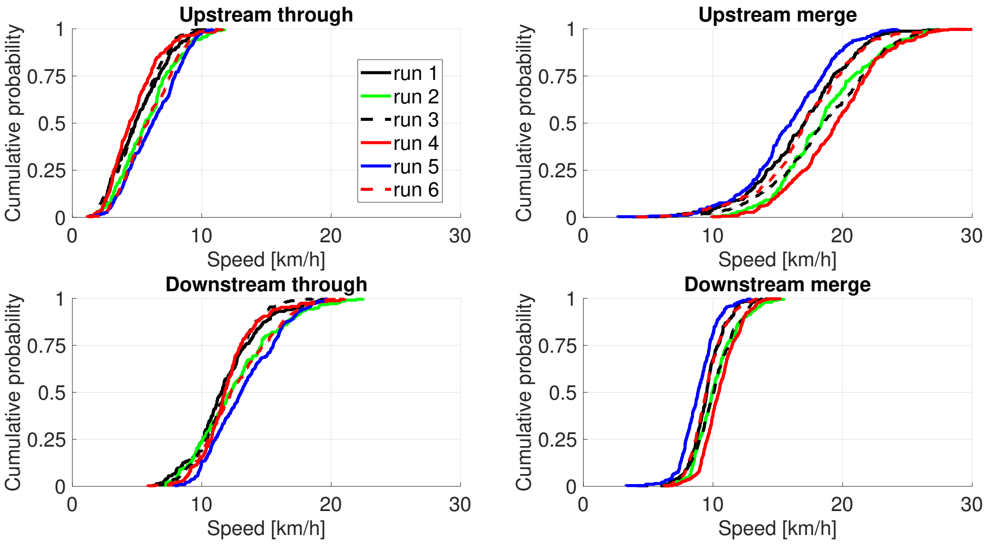


Figure 4.3: Distribution of average speed per run, section and direction.

The different cycling speeds affect travel time, whose average values per run and direction are shown in Table 4.2, along with the corresponding standard deviation. It can be seen that the variance is higher for through cyclists, which can be explained by the dependence of travel time on the number and duration of the stops to give way to the merging cyclists. Moreover, through cyclists need on average twice the time to pass, which is attributed partially to the longer distance in the camera view and partially to the delay from the stops.

Table 4.2: Travel time per run and direction.

Run	Merging cyclists		Through cyclists	
	Average [s]	Std. [s]	Average [s]	Std. [s]
1	3.20	0.87	6.61	4.21
2	3.47	0.52	5.40	2.42
3	3.18	0.81	6.65	3.43
4	3.29	2.83	6.59	4.04
5	3.63	1.04	5.18	2.73
6	3.42	0.84	6.09	4.03

Levene's test is performed to test for equality of variances in travel time (Field 2013). The outcome is that there is homogeneity of the travel time variances of merging cyclists, and heterogeneity of variances for through cyclists. This means that an ANOVA needs to be applied to compare the average travel times of merging cyclists, and the Kruskal-Wallis test needs to be performed on the travel times of through cyclists (Field 2013). Both tests result in a p-value $< 0.01\%$, which indicates that the difference is statistically significant between the runs for both directions.

For further insights, KS-tests are performed to compare the underlying travel time distributions. According to the results of the tests, the travel time distributions of through cyclists are the same for runs 1, 3 and 4. This means that there is no learning effect, and that the presence of different bicycle types does not appear to affect the travel time. Moreover, the lane marking in run 6 leads to travel time savings compared to run 4 given the same cyclist population and the aforementioned lack of learning effect. However, the travel time distribution of run 6 is not statistically different from those of runs 2 and 5. This proves that the heterogeneity of the cyclist population is at least as important in decreasing travel time as the presence of lane marking. Regarding merging cyclists, runs 2, 3 and 4 are the fastest, and their travel time distributions are not statistically different. Runs 1 and 6 follow the same travel time distribution and are slower than 2, 3 and 4. The fact that run 3 is faster than run 1 confirms that there is a learning effect for merging cyclists which makes them faster. The lane marking, despite this learning effect, slows them down (run 6 is slower than 4). As runs 2 and 4 follow the same distribution, there is no evidence for an effect of bicycle type on travel time. The heterogeneity stemming from different personal characteristics is, on the other hand, shown to have a positive effect on travel time. This is because of the difference found in the distributions of runs 1 and 2, but also of runs 4 and 5. To conclude, the lane marking is overall less effective in decreasing the travel time compared to familiarity with a situation (i.e. learning effect) and cyclist heterogeneity.

4.5.2 Efficiency based on infrastructure use

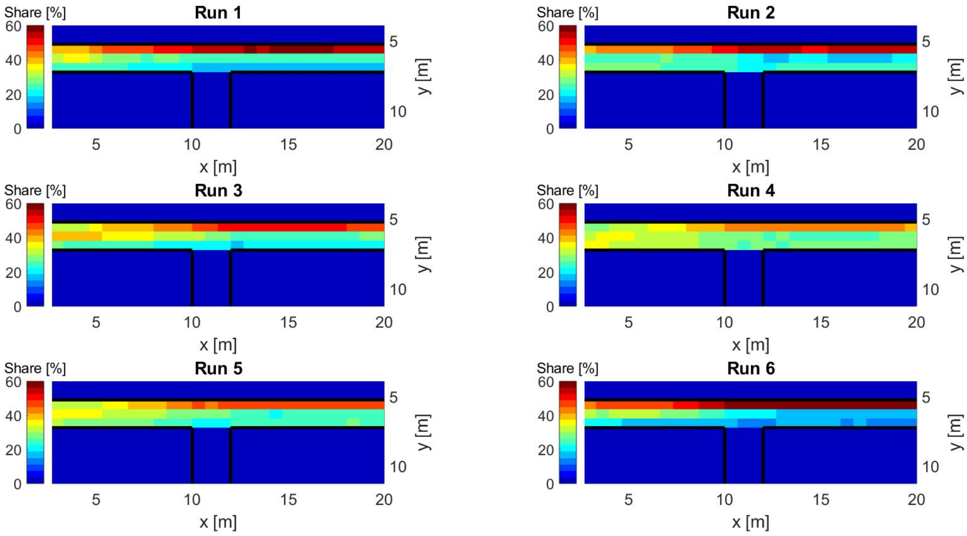
The space utilisation is visualised per run and direction in the heatmaps in Figure 4.4. The colour indicates the share of cyclists that rode over each square of the infrastructure. Note that the maximum share value is different for the two directions; it reaches 60% for through cyclists and 80% for merging cyclists. This means that through cyclists make better use of the full width of the cycle path, while merging cyclists are more concentrated in an optimal trajectory.

Figure 4.4(a) shows the heatmap of the infrastructure usage by through cyclists. It is remarkable that the sublane closest to the left edge of the infrastructure is used by more than half of the cyclists, especially just upstream of the merging location and downstream till the end of the camera view. An exception to this observation is run 4, where the entire width of the cycle path is equally used by through cyclists. In general, the utilisation of the right and middle sublanes seems greatly dependent on the run and group of cyclists. In some runs they are equally likely to be used by through cyclists, while in others the middle sublane is more prominent. The section upstream has the highest variability in use, as some cyclists might not mind to yield and others might shift to the left sublane hoping to avoid a stop. In the case of run 6, the lane marking causes a very small spread throughout the path, as cyclists are concentrated in the left lane both upstream and downstream. It is particularly interesting to notice the difference from run 4, which has the same population and where through cyclists seem to claim the whole width for their own use. This means that with the addition of the lane marking, most cyclists are already in the left lane upstream of the T-junction as intended by the marking. The lane marking, thus, counters the learning effect observed when comparing runs 1 and 3. These runs show that as cyclists get more familiar with the infrastructure (run 3), they optimise its usage by spreading more equally over the three sublanes to improve efficiency. The presence of the lane marking prevents them from doing so and pushes them even more into the left sublane.

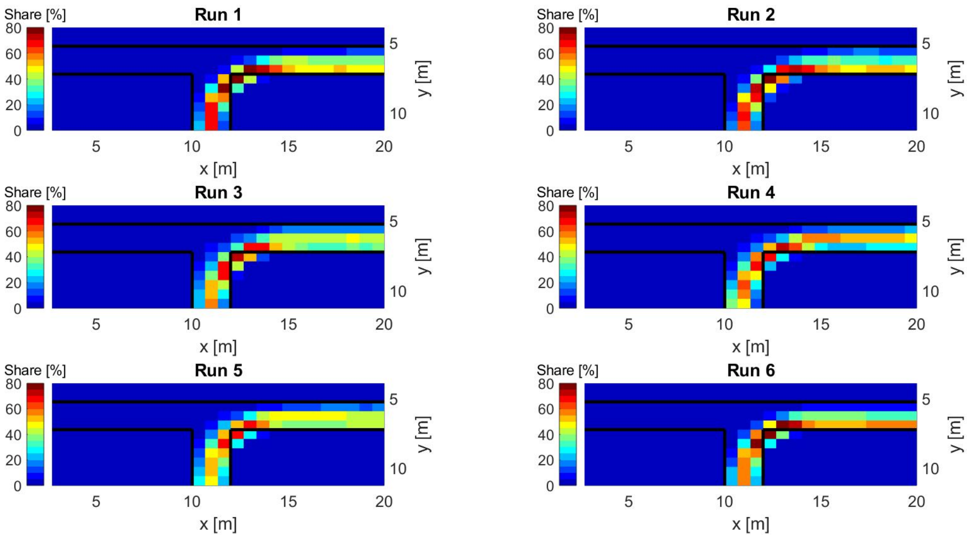
The concentration of through cyclists on the left lane in run 6, gives space to the merging cyclists to use the right sublane, as proves to be the case in Figure 4.4(b). This figure displays the patterns of infrastructure use by the merging cyclists. It can be seen that in all runs, most merging cyclists use the middle and right sublanes downstream of the T-junction. In most cases the share of cyclists in these two sublanes is equal, with a slight preference for the right sublane. The exception is, again, the cyclist population of run 4, where the middle sublane is used by half of the merging cyclists. For these cyclists, the lane marking pushes them to the right of the cycle path and leads to a utilisation of the infrastructure that other cyclist populations could achieve without the lane marking. Regarding the use of the upstream section by merging cyclists, more than half of them use the middle sublane and start shifting to the right sublane at about 2 m upstream of the T-junction.

As already mentioned, merging cyclists seem to follow an “optimal” trajectory. They start in the middle sublane and move to the right to take the turn. As expected, the right angle is too sharp and most cyclists violate the track boundary at the junction. A curved edge that would permit passage into the 67 x67 cm^2 would be optimal for most merging cyclists. In order to accommodate all cyclists, the turn should have a radius of 2 m. This value could actually be smaller, since the cyclists lean inwards when taking the turn and their heads are being tracked through the red caps for the construction of the trajectory. The observed trajectory is, therefore, shifted slightly to the right compared to where the bicycle is. However, the effect is considered negligible due to the low cycling speeds when taking the turn, and is ignored.

Figure 4.5 displays the cumulative probability headway distribution functions per run. Kolmogorov-Smirnov tests indicate that only the distribution of the first run has a statistically significant difference from the rest. The explanation for this might be the learning effect, as this is the first run during which the participants could make use of the T-junction and approached more cautiously than later in the experiment. Another explanation might be related to the demand (i.e. the offered gaps), as one over the mean headway is equal to the average flow of one cyclist per second. In either case, this result shows that the lane marking does not have an influence on the observed headways and, therefore, it is unlikely that it influences the capacity of the junction. Estimation of composite headway models will be performed in the future to provide further evidence.



(a) Through cyclists.



(b) Merging cyclists.

Figure 4.4: Space utilisation per run and direction. Note that the scale of the colour bar is different for the two directions.

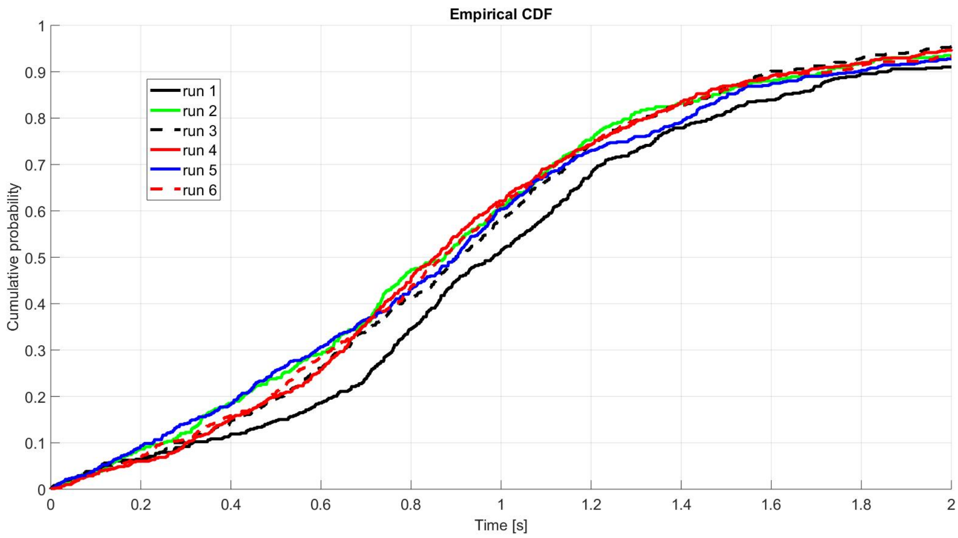
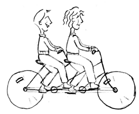


Figure 4.5: Headway distribution per run downstream of the T-junction.

4.5.3 Efficiency based on interaction with other cyclists

Table 4.3 shows the number of through cyclists that are observed per run in relation to the interacting merging *group size*. The merging group size is the number of merging cyclists upstream of the T-junction when a through cyclist passes the 6 m cross section. Therefore, a group size of 0 means that there are no merging cyclists upstream of the T-junction when the through cyclist is at the 6 m cross-section, and thus that the considered through cyclist does not have an interaction. The total of each row in the table is the total number of through cyclists in the corresponding run. As the duration of the runs differs (see Table 4.1), the absolute number of through cyclists is a function of the corresponding run duration. For this reason, the totals are normalised per minute of run duration. The absolute numbers are shown on the left part of the table and the normalised ones on the right. As an example, in run 3, it occurred 40 times that a through cyclist was interacting with 2 merging cyclists when crossing the 6m cross section. The duration of run 3 was 10 minutes, so on average there were 4 through cyclists per minute interacting with 2 merging cyclists when crossing the 6m cross section.



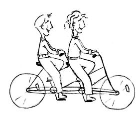
The primary aim of this table is to demonstrate that most through cyclists have either no interaction or interact with one merging cyclist. Moreover, it is noted that the maximum merging group size is 4 cyclists (though it occurs only once) and that the number of observations decreases as the merging group size increases. A Chi-square test is performed on the normalised results, which shows that there is no statistically significant difference between them.

Table 4.3: Number of through cyclists per run and interacting merging group size.

Run	Total through cyclists					Average through cyclists per minute				
	Merging cyclist group size					Merging cyclist group size				
	0	1	2	3	4	0	1	2	3	4
1	106	104	23	2	0	12	12	3	0	0
2	112	80	30	0	0	17	12	5	0	0
3	107	124	40	5	0	11	12	4	1	0
4	126	97	34	7	0	12	10	3	1	0
5	165	113	43	12	1	17	11	4	1	0
6	149	160	67	10	0	11	12	5	1	0

In order to get a clearer picture of the groups that interact, the distribution of these totals over different sizes of through cyclists is needed. The group size for through cyclists is calculated when each through cyclist crosses the 6 m cross section. The group size is equal to the number of cyclists upstream of the T-junction at that moment. Table 4.4 summarises this distribution for each run. The values in each cell correspond to the share of through cyclists that interacted with a particular merging group size. So 0% means that the specific combination of through and merging group sizes was not observed during that run, while for example in run 3 (subtable (c)), there are in total 40 through cyclists that interacted with a group size of 2 merging cyclists and for 5% of these cyclists (i.e. 2 times) there were 7 through cyclists upstream of the T-junction, including the through cyclist that is crossing at that moment the 6 m cross section.

The overall maximum group size of through cyclists observed is 10 cyclists, which means that the maximum expected size of through cyclists is observed but not that of merging cyclists. This can be explained by the fact that through cyclists stop to give way, and therefore, form



a queue upstream of the junction. The most common situation in all runs is to have one merging cyclist and up to three through cyclists. The effect of the lane marking in run 6 is that longer queues of through cyclists are observed.

Table 4.4: Distribution of through cyclists interactions per run (subtable), group size of merging cyclists (column) and group size of through cyclists (rows).

(a) Run 1

	1	2	3
1	19%	22%	50%
2	19%	13%	0%
3	35%	43%	50%
4	13%	22%	0%
5	11%	0%	0%
6	3%	0%	0%
7	0%	0%	0%
8	0%	0%	0%
9	0%	0%	0%
10	0%	0%	0%
Total	104	23	2

(b) Run 2

	1	2	3
1	24%	16.6%	0%
2	31%	20%	0%
3	27.5%	36.7%	0%
4	12.5%	26.6%	0%
5	4%	0%	0%
6	1%	0%	0%
7	0%	0%	0%
8	0%	0%	0%
9	0%	0%	0%
10	0%	0%	0%
Total	80	30	0

(c) Run 3

	1	2	3
1	16%	12.5%	20%
2	34%	17.5%	0%
3	24%	35%	80%
4	10.5%	10%	0%
5	10.5%	7.5%	0%
6	2%	12.5%	0%
7	2%	5%	0%
8	1%	0%	0%
9	0%	0%	0%
10	0%	0%	0%
Total	124	40	5

(d) Run 4

	1	2	3
1	20%	21%	0%
2	29%	26%	43%
3	27%	32%	43%
4	14%	12%	14%
5	7%	9%	0%
6	3%	0%	0%
7	0%	0%	0%
8	0%	0%	0%
9	0%	0%	0%
10	0%	0%	0%
Total	97	34	7

(e) Run 5

	1	2	3
1	16%	33%	17%
2	40%	28%	42%
3	25.5%	9%	25%
4	10.5%	23%	0%
5	6%	2%	8%
6	1%	0%	0%
7	1%	5%	8%
8	0%	0%	0%
9	0%	0%	0%
10	0%	0%	0%
Total	113	43	12

(f) Run 6

	1	2	3
1	18%	19.4%	10%
2	29%	28.4%	20%
3	21%	28.4%	40%
4	17%	16.4%	0%
5	9%	3%	0%
6	4%	3%	0%
7	0%	1.4%	0%
8	1%	0%	10%
9	0%	0%	20%
10	1%	0%	0%
Total	160	67	10

Table 4.5 summarises the properties of the stops made in each run. The first column shows the run number. Columns two and three display the average and standard deviation of the stopping position in the longitudinal (i.e. x) direction which is the direction of movement. The next two columns refer to the average stopping position in the lateral (i.e. y) direction, averaged separately over each lane. Figure 4.2(c) has

already introduced the basic cross-sections of the cycle path which are crucial in interpreting the values of Table 4.5. More specifically, it is known that the longitudinal distance at which the through cyclists enter the conflict area of the T-junction is at 10 m, while in the lateral direction the right lane is between 6.5 m and 7.5 m, and the left lane between 5.5 m and 6.5 m. The sixth column shows the share of the stops that take place in the left lane. Column seven and eight refer to the average and standard deviation of the stop duration, respectively. The last column displays the stopping frequency S expressed in stops per minute.

The stopping positions in the longitudinal direction are tested with Levene's test and found to meet the homoscedasticity assumption. Only cyclists stopping at the front of the queue are considered in the analysis of the longitudinal stopping positions. ANOVA shows that there is no statistically significant difference at a 5% significance level. The average stopping position is, thus, at around 8.7 m, which justifies the choice of the 6 m cross-section as the decision point for the reaction to the encounter.

In the lateral direction, the two lanes that are formed by the lane marking are separately considered. All stopping cyclists are taken into account in the analysis. The stops in the right lane are at 7 m and at 6 m in the left lane, which are the middle points of the width of the corresponding lane. Though the difference in these values is not statistically significant between the runs, the share of cyclists that stop in each lane changes. This share ranges from 30% to 51%. The highest value is observed in run 6, but it is very close to the 48% of run 4, which has the same cyclist population. The lane marking does, thus, not affect the stopping position. The cyclist population plays a more important role in this decision.

It is remarkable to note that only through cyclists stopped to give priority and no merging cyclist. Based on this, it could be concluded that the priority perception is not inverted at bicycle T-junctions, though one should be aware of the experimental conditions.

The differences in the *stop duration* are statistically tested. Levene's test rejects the assumption of homoscedasticity. The applied Kruskal-Wallis test leads to the conclusion that at a 5% significance level the medians of the stop duration of the six runs have a statistically significant difference. In this case, the introduction of the lane marking makes the stops of the through cyclists last longer, which means that it reduces the efficiency of the T-junction. A possible explanation for the longer stops is that the group sizes that interact are larger and the queues that are formed are longer, so the cyclists that are at the back of the queue need to wait a long time.

The *stopping frequency* (number of stops per minute) varies a lot among the runs. The averages, however, should not be used for the comparison as they are sensitive to outliers and Levene's test rejects the hypothesis of equal variances. Therefore, the Kruskal-Wallis test is performed. The outcome is that there is no statistically significant difference in the medians of the stopping frequency of the six runs at a 5% significance level. Apparently, the lane marking does not have a major influence on the number of stops. This means that despite the lane marking, cyclists felt obliged to stop just as frequently and give priority to merging cyclists who would not be able to merge into the main track comfortably otherwise.

Table 4.5: Properties of yielding stops.

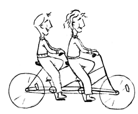
Run	Position in x [m]		Position in y [m]		% stops in	Duration [s]		Frequency S
	Average	Std.	Right lane	Left lane	left lane	Average	Std.	[stops/min]
1	8.4	1.4	6.9	6.0	45	6.0	3.8	46
2	8.5	0.7	7.0	6.2	30	4.6	2.2	7
3	8.7	1.4	6.9	5.9	44	4.5	2.5	77
4	8.6	0.7	6.8	5.9	48	3.5	2.1	54
5	9.0	0.8	6.9	6.0	37	4.4	2.2	50
6	8.8	0.7	7.0	5.9	51	6.5	4.7	180

4.6 Conclusions and recommendations

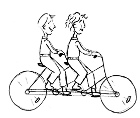
We investigated the cyclist behaviour at T-junctions dedicated to bicycles and assessed the efficiency of the T-junction design on the bicycle flow. The analyses were performed using empirical trajectory data collected during a large-scale cycling experiment. The effect was studied of adding lane marking that advised through cyclists to shift to the left so that merging cyclists could occupy the space to the right of the cycle path. The expected outcome was that the two conflicting streams would be separated in space, thereby increasing flow efficiency. In order to assess the efficiency, a framework was proposed with eight indicators that cover different aspects of the infrastructure efficiency.

The findings overall suggest minimal to no effect on the efficiency once the lane marking is introduced, but great effect resulting from the heterogeneity of the cycling population, especially stemming from personal characteristics and less so from the different bicycle types. The speeds of through cyclists increased by adding the lane marking, and not by the repetitive nature of the experiment. However, cyclists in other runs could reach even higher speeds, indicating that the personal characteristics are at least equally important. With respect to merging cyclists, it was shown that there is a learning effect towards an "optimal" trajectory, which had a positive effect on speed. Despite the increased confidence due to familiarity, the introduction of the lane marking slowed merging cyclists down, as it made the perception of the 90 degree turn even sharper, forcing them to use only the space of the right lane. This way, the desired outcome of clear lane separation was achieved, while at the same time it was observed that cyclists in some runs could self-organise, without guidance, into such separated flow. Regarding the capacity, no change was observed in the time headways. The same can be stated about the stopping frequency of through cyclists, even though the duration of the stops increased and longer queues were formed. The combination of these findings means that there is no obvious advantage, as originally hypothesised, in the introduction of the lane marking, but possibly a negative one, as the duration of the stops, and hence the delay of the through cyclists, increased.

Based on these findings, the design recommendation would be against using such a lane marking at bicycle T-junctions. Through cyclists are capable of making space to allow merging cyclists to fit in the cycle path without instructions. This self-organisation might even prove more efficient as they can use the full width of the cycle path when no encounter is about to take place. Moreover, the trajectories of merging cyclists showed that a turn with a radius of 2 m would accommodate all cyclists. In CROW (2016), a minimum radius of 5 m is advised, such that the speed of cyclists is not interrupted. It is indeed the case that the speed is reduced, but this more cautious approach might lead to safer interactions at the T-junction.



An important outcome of our analysis is that no merging cyclist stopped to give priority to through cyclists. This could indicate that cyclists respect the priority traffic rules and there is no inverted perception of priority at bicycle T-junctions. An alternative explanation for this might be the experimental setting under which participants felt obliged to abide by the traffic rules and yield to traffic coming from the right. In order to conclude on perceived priority, real-world observations should be collected and analysed. Moreover, the effect of a wider range of densities on the cycle path should be investigated, along with the effect of different cycle path widths. Additionally, a topic for future research is to compare the findings of this study with the efficiency at a T-junction with priority given to the through traffic, either by means of signage or in countries where traffic from the left has priority and based on that conclude which rule leads to the highest efficiency. Apart from this, it is of interest to investigate when a decision to yield is made instead of a decision to make space. Modelling this decision making process will be the focus of future research.

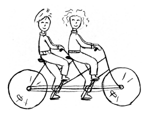


Chapter 5

Cycling behaviour at unsignalised intersections

In the previous chapter, the cycling behaviour and bicycle flow efficiency at a T-junction was investigated using the trajectory data from the large-scale, controlled cycling experiment presented in chapter 3. This chapter goes a step forward from the data analyses into the mathematical modelling of the behaviour based on this rich dataset. The focus of this chapter is on a four-legged bicycle intersection. For the mathematical modelling, the two-layer modelling framework introduced in chapter 2 is revisited, updated and used to capture the yielding behaviour of cyclists, both in terms of decision making and movements, upstream of an unsignalised bicycle crossing. The estimated yielding models and corresponding behavioural insights are the main contributions of this chapter. What is more, through this application, the applicability and extension towards generalisability of the two-layer framework is demonstrated, thereby establishing it as a viable modelling paradigm to capture operational cycling behaviour.

This chapter is based on the submitted article: Gavriilidou, A., W. Daamen, Y. Yuan, S. P. Hoogendoorn (2020a) To yield or not to yield? A behavioural model at unsignalised bicycle crossings, submitted to Transportation Research Part C: Emerging Technologies.



5.1 Introduction

Intersections are critical locations in a bicycle network as the probability that a cyclist is involved in an accident is higher (Dozza & Werneke 2014; Flower & Parkin 2019). This higher risk is attributed to the infrastructure design (e.g. turns, road marking, lighting conditions) (Wijlhuizen et al. 2016), but also to the number of interactions that take place. Imbert & te Brömmelstroet (2014) showed that on crowded cycle paths at an intersection, cyclists experience a higher level of stress and discomfort, which makes them adhere less to traffic rules, thereby increasing the risk of collisions. The question is whether intersections dedicated to bicycles, where only bicycle-to-bicycle interactions take place, elicit such unsafe behaviour as well. Answering it gives insights into safer intersection design. Moreover, understanding the decision making process of intersecting cyclists and their movements upstream of the intersection may reveal factors that affect the collision risk.

After investigating the behaviour at T-junctions, we found that cyclists adhere to traffic rules (i.e. give priority to cyclists coming from the right), while self-organising within the cycle path to make efficient use of the available space (Gavriilidou et al. 2020b). In this context, self-organisation refers to lane changing manoeuvres by non-priority cyclists such that they can continue without stopping on the left side of the cycle path and merging cyclists from the right can enter on the right side. These manoeuvres were more efficient when left to the judgement of the cyclist rather than when indicated by the infrastructure. As the empirical findings at the bicycle T-junction did not show evidence of unsafe interactions, we focus on a four-legged bicycle intersection for further empirical insights.

In addition to these insights, we develop a cyclist yielding model. Existing literature has investigated the trajectory deviation of bicycles in conflict with other modes at unsignalised intersections (Huang & Wu 2009; Huang et al. 2017; Zhang et al. 2017), but not the decisions related to priority allocation. Moreover, the yielding behaviour of motorised traffic when interacting with bicycle traffic has been researched (Hydén et al. 2007; Phillips et al. 2011; Svensson & Pauna-Gren 2015; Ng et al. 2017; van Haperen et al. 2018). However, to the best of the authors' knowledge no study has previously investigated the yielding behaviour of crossing cyclists. The term yielding is used for situations when two

cyclists are on collision course. The cyclist who lets the other cross first is the one yielding (Svensson & Pauna-Gren 2015).

In this chapter, we use empirical trajectory data collected during the large-scale cycling experiment presented in chapter 3 and build upon the two-layer modelling framework developed in chapter 2 to capture the mental and physical processes of operational cycling yielding behaviour. The empirical findings and behavioural insights generated in this research are thus one of its main contributions. Another contribution is the estimated discrete choice models for both layers that describe the cyclist decision making and movements upstream of an unsignalised bicycle crossing. These models predict cycling behaviour when implemented in a simulation model. Last but not least, through this application, the use and extension towards generalisability of the two-layer framework is demonstrated, thereby establishing it as a viable modelling paradigm to capture operational cycling behaviour.

The chapter is structured as follows. In section 5.2 the two-layer modelling framework is revisited and extended, as well as elaborated for the case of yielding behaviour. Section 5.3 presents the data at the crossing from the controlled experiment. The model estimation approach for each layer is discussed in section 5.4. In section 5.5 the results of the best performing estimated model for each layer are presented, along with simulation results for face validation. Finally, in section 5.6 conclusions are drawn and recommendations for future research are made.

5.2 Improving two-layer modelling framework

The two-layer framework to model operational cycling behaviour is shown in Figure 5.1 taken from (Gavriilidou et al. 2019a) and adjusted for the application to the yielding situation at unsignalised intersections. The attributes in the box linked to the yielding behaviour are the result of a stated preference survey, discussed in (Gavriilidou et al. 2019b). The ones in black are those investigated in this research.

According to this framework, a choice whether to yield or not is first made in the mental layer and is then communicated into the physical layer. The physical layer describes the choices cyclists make in consecutive time steps regarding changes in their direction (steering) and speed (pedalling). The new states of the cyclists that emerge from the physical layer feed back into the mental layer to allow for decision updates when necessary. An extension of the existing framework are the two arrows within the mental layer pointing towards the queuing position. When the output of the mental layer choice for yielding is to come to a complete stop, then a queuing position decision needs to be made before proceeding with decisions in the physical layer. This is represented by the black arrow in the mental layer going from yielding to queuing position. The same process holds when a decision is made to stop at a red light over a red-light-running decision (grey arrow from stopping at red light to queuing position). In this research, we investigate the yielding model and corresponding physical layer, as well as the black arrows going in and out of these boxes.

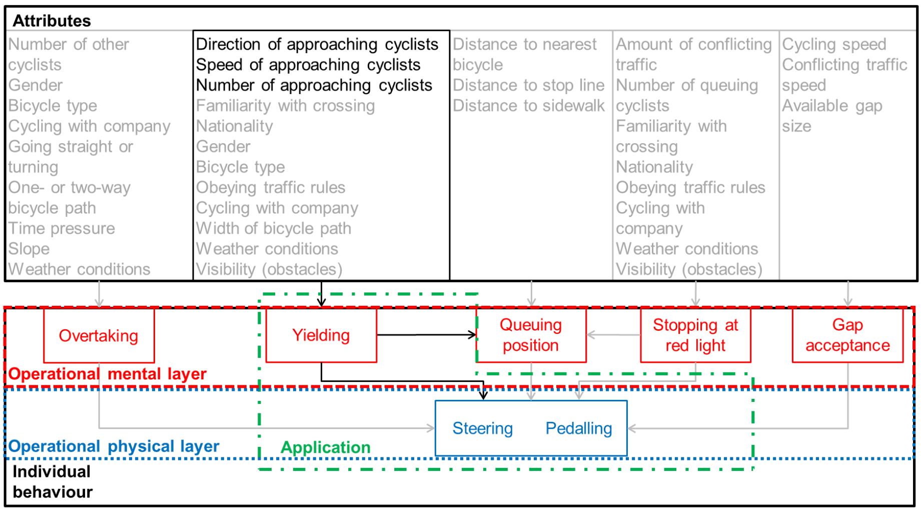


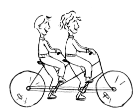
Figure 5.1: Improved two-layer framework, adjusted from Gavriilidou et al. (2019a). The elements that go beyond the scope of this chapter are coloured in grey, while the scope of the model application in this chapter is framed within the green box.

In order to model the decision making upstream of an unsignalised bicycle crossing, it needs to be clear which process in the mental layer is active. To facilitate this, as an additional contribution to the improved framework, we assume that approaching cyclists follow the decision process displayed in Figure 5.2. If the cycle path is blocked by other cyclists who decided to stop and give priority to the other cyclist stream, then the approaching cyclist needs to queue. Therefore, a mental layer decision for a queuing position is assumed to be made instead of a yielding decision. In reality it might be possible for the cyclist to go out of the cycle path and create a crossing opportunity but we consider the chances of this negligible and exclude these cyclists from the yielding decision model.

When there is sufficient space in the cycle path for the approaching cyclist to reach the intersection, a yielding decision must be made in the mental layer. The parameter set that determines the output of the yielding decision is hypothesised to differ for three cycling groups, which is visualised by the three arrows leading into the yielding box. The first group consists of cyclists who have priority based on traffic rules. These are referred to as major stream cyclists in the ensuing. They are expected to be more aggressive in their crossing decision and not willing to yield to cyclists approaching from the non-priority direction. The second and third groups are called minor stream cyclists; they are cyclists who do not have priority according to traffic rules. Thus, they are expected to react and yield to cyclists coming from the priority direction.

The difference between the two groups lies in the presence of conflict with major stream cyclists upon arrival at the crossing if the minor stream cyclists would continue with their cycling speed at the moment the yielding decision is made. If no conflict is anticipated (group 2), then the behaviour of the minor stream cyclist is expected to be less cautious. When a conflict is anticipated (group 3), the minor stream cyclist needs to decide whether or not to yield to the major stream cyclist. A yielding decision is the choice to give priority to the major stream cyclist by reducing speed and/or changing lane. In some cases, the speed reduction can be such that the yielding decision is a stopping decision. For those cyclists, a queue position must be decided prior to determining their trajectory in the physical layer. So their next decision is still in the mental layer, and specifically in the queuing position choice model. All other yielding (and non-yielding) decisions contain sufficient information for the physical layer to be activated and determine the cyclist trajectory.

Regarding the mathematical modelling, discrete choice theory and the utility maximisation principle are used, similar to Gavriilidou et al. (2019a), as the behavioural assumptions remain the same.



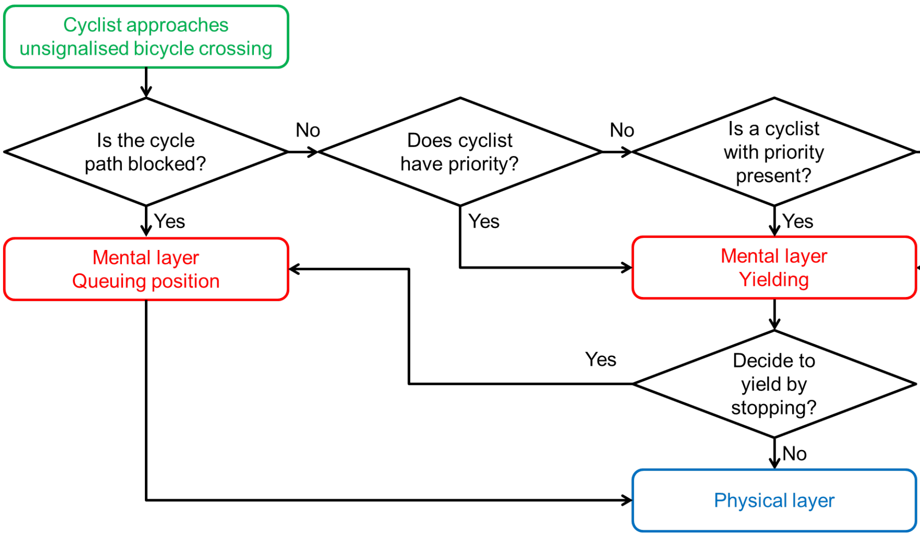


Figure 5.2: Decision flow chart for a cyclist that approaches an unsignalised bicycle crossing.

In the operational mental layer of this application, cyclists need to decide the moment they will arrive at the crossing, as well as their lateral position within the path when they arrive. The longitudinal position is known and fixed for all cyclists; it is the cross-section just upstream of the intersection itself for the corresponding cycling stream. This cross-section is referred to as cross line in the ensuing. In the operational mental layer model, the space is discretised into lanes and time is discretised in comprehensive time intervals that depend on the clearance time of the intersection. The motivation of the specific grid is explained in section 5.4.1. The different combinations of arrival time and sublane are the alternatives that compose the choice set. Each alternative is assigned a (dis)utility based on attributes of the alternative and characteristics of the cyclist. Availability conditions are also taken into account, since cyclists should not be allowed to arrive at the crossing at the same time and sublane to avoid collisions.

If the output of the mental layer model does not correspond to a stopping decision, then the selected arrival time-sublane combination is fed as input to the discrete choice model of the operational physical layer, where the cyclist decides upon the controls to reach the desired position within the desired time. As the operational physical layer is, according to the framework, generic to all situations, its principle remains

the same as in (Gavriilidou et al. 2019a). Cyclists choose a combination of pedalling and steering, which are expressed as changes in speed and direction relative to the speed and direction at the moment the decision is made.

Figure 5.3 demonstrates the operational yielding model for both layers. In the mental layer the arrival time and sublane (indicated in red) are chosen and provided as input to physical layer. In the latter, a sequence of decisions (blue cells) is made that corresponds to the combination of angle and speed difference with the highest utility at each time step (grey-scale fans, where a lighter shade of grey is given to the grid for decisions to be made in further in the future compared to the current time step). The sequence of positions resulting from these choices forms the cyclist trajectory.

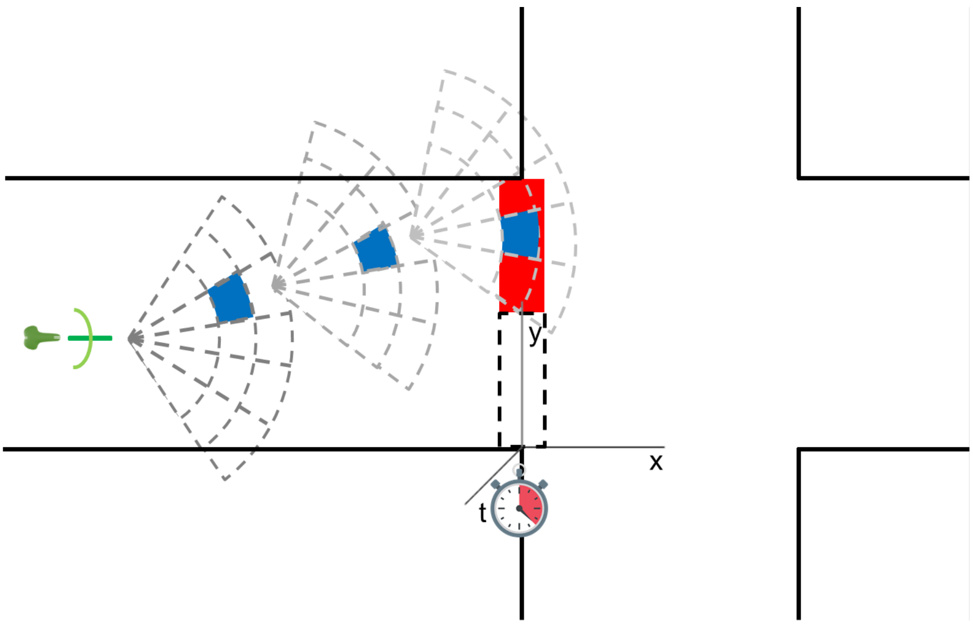


Figure 5.3: Schematic of the operational yielding model (mental layer decisions in red and physical layer decisions in blue) at an unsignalised bicycle crossing. The grey-scale fans of the physical layer correspond to different decision moments throughout the trajectory.

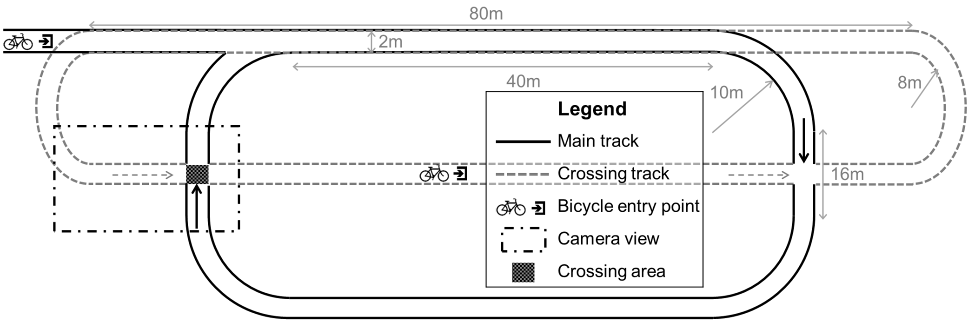
5.3 Data at unsignalised bicycle crossing

The dataset used to estimate and validate the cyclist yielding behaviour model consists of cyclist trajectories collected during the controlled experiment presented in chapter 3. In this section, details about the experimental runs with intersecting bicycle flows are provided (subsection 5.3.1), followed by the description of the data processing to prepare the dataset to estimate the models in this chapter (subsection 5.3.2). In the last subsection (5.3.3), the classification of the cyclists in the collected dataset is presented. This classification divides the cyclists in the three groups whose behaviour has been hypothesised to differ (Figure 5.2). For more information about the experiment and the demographics of the participants the reader is redirected to Gavriilidou et al. (2019b).

5.3.1 Crossing scenarios

During the controlled experiment, three runs of ten minutes each took place on the track layout shown in Figure 5.4(a). These runs were spread throughout the day to keep the schedule of the participants interesting, alternating between different routes and interactions. The track was marked by white tape and was 2 m wide at all locations. The participants were randomly divided into two groups. Each group was instructed to follow a particular route for the duration of the scenario. In order to enter the track, each group used its own entry point as indicated in the layout. Cyclists assigned to the main track (full black line) entered at the top left, while those following the crossing track (grey dashed line) entered in the middle area of the main track. The track was always marked with a continuous tape, so it is only for readability of this figure that the crossing track is indicated with a dashed line.

This layout creates two intersection points for these two bicycle streams, one on the left and one on the right, and a bidirectional flow on the top part of the main track. The split of the participants over the two tracks was 65-35%, with the majority following the main track. This means that most conflicts that require a yielding decision by the minor stream take place at the crossing on the left indicated as "crossing area" in Figure 5.4(a). For this reason, in this chapter, we use the cyclist trajectories extracted from the half hour of video data from the overhead camera placed at the left crossing. A snapshot of the view from this camera is shown in Figure 5.4(b). The major stream cyclists go from bottom to top, and the minor stream from left to right. In this snapshot, five minor stream cyclists have stopped to give priority to a major stream cyclist and one of them is restarting, entering the crossing area while the major stream cyclist clears the crossing area.



(a) Layout of experimental track.

(b) Snapshot at left crossing.

Figure 5.4: Experimental track layout and camera snapshot at the left crossing.

Regarding the cycling behaviour, participants were instructed to cycle as they would normally do, and to try to stay within the edges of the cycle path as much as possible. The given instructions left the choice of speed, priority allocation and overtaking to the participants. The tape was thick enough to let passing cyclists know that they are crossing it, but not raised as a curb that could cause safety incidents.

5.3.2 Data preparation

The cyclist trajectories have been extracted by tracking the position of the red caps, worn by all participants, in each frame of the videos (every 0.04 s), orthorectified and cleaned, following the steps described by Gavriilidou et al. (2020b).

The extracted trajectories are further smoothed using a rolling horizon of 0.5 s to get rid of outliers related to the detection accuracy. This processing, thus, makes the speed profiles more reliable and suitable for use in the model estimation.

For each cyclist, only the part of the trajectory that is upstream the crossing and up to the cross line is relevant and kept in the dataset. Moreover, for the minor stream cyclists the first four meters in the camera view are discarded. As seen in Figure 5.4, minor stream cyclists follow a curve of 8 m radius and then a straight stretch of 8 m to reach the crossing. Part of the curve is visible in the camera view and thus also in the upstream part of the trajectories. In order to minimise the effect of the curve on the observed steering behaviour, the first part is removed from further analyses.

Last but not least, cyclists who encounter a blocked path ahead of them are removed from the dataset. A blocked path corresponds to at least one cyclist stopped in each sublane of the cycle path upstream of the crossing area, thus not allowing any space for new arrivals to cross. As these cyclists are forced to stop and join the queue, they choose a queuing position rather than make a yielding decision (Figure 5.2).

5.3.3 Cyclist classification

According to the decision flow chart of Figure 5.2, the yielding behaviour of cyclists is expected to differ depending on whether they have priority or a conflict with someone who has priority. It is, therefore, important to identify which cyclists in the dataset belong to which of these groups.

The classification into major stream and minor stream is very simple as it follows the traffic rules stating that cyclists coming from the right have priority. All cyclists who ride from bottom to top in the camera view are, thus, major stream cyclists, and those who ride from left to right are minor stream cyclists. In the collected dataset, there are 938 trajectories in the major and 315 in the minor stream.

In order to determine whether a minor stream cyclist has a conflict with a major stream cyclist, it is checked whether the crossing area is occupied by a major stream cyclist. The moment of interest for this check is the expected arrival time at the cross line if the decision maker would maintain the cycling speed at the moment the decision is made. So, a zero-acceleration strategy is used as the rule defining a conflict. The motivation behind this choice of strategy is that it is in line with the least effort principle that cyclists have been assumed to follow. If by making no speed changes the cyclist would collide, then there is a conflict and, for the collision to be prevented, at least one of the cyclists needs to react. By introducing this zero-acceleration strategy, it is possible that some cyclists are misclassified. The extent of this misclassification can be checked in the model results.

The quality of this classification depends on the accuracy of the anticipation by minor stream cyclists of the movements of major stream cyclists. In reality, a prediction is made based on assumptions about the speed and acceleration of the conflicting cyclist. For the model estimation, the exact passing moments of the major stream cyclists are used, which are known from the collected dataset. As a result, out of the 315 minor stream cyclists, 104 are categorised to be in a conflict when following a zero-acceleration strategy.

Using the exact passing moments corresponds to the ideal, highest accuracy, anticipation. This is most likely different from what the cyclists consider when making a decision, which makes it hard to assess the effects of perfect anticipation. At the same time, it is not yet known what horizon cyclists consider (i.e. how far ahead they look) and what the accuracy of their anticipation is. Future research should look into these matters.

To sum up, following the aforementioned classification the dataset contains the decisions of 938 cyclists in the major stream class, 211 in the minor stream class without conflict and 104 in the minor stream class with conflict.

5.4 Yielding model estimation approach

In this section the estimation approach is discussed for each layer separately (subsection 5.4.1 for the mental layer and subsection 5.4.2 for the physical layer). The approach covers the justification of the choice set per layer and the selection of attributes to explain the corresponding behaviour, as well as assumptions specific to the model estimation.

5.4.1 Operational mental layer estimation approach

In the mental layer model of the yielding choice, cyclists select their arrival time at the crossing as well as their lateral position in the path at the cross line. This choice is made when they exit the curve and have the crossing in view. In this application, we assume that the choice is made only once, but future research could look into updating over time. The space is discretised into two sublanes, such that the chosen lateral position is either in the left or in the right sublane. The choice for two sublanes is motivated by the resulting sublane width, which is 1 m and is sufficient to fit the 70 cm handlebar width (CROW 2016) along with a safety margin that is needed because the cyclists are moving at normal cycling speeds. Moreover, during the experiment two sublanes were observed to be formed on the track. Another confirmation for the use of two sublanes comes from the observed lateral positions of cyclists when arriving at the cross line. This can be seen in Figure 5.5, where more than 99% of the cyclists have a lateral position (i.e. distance from the left edge of the cycle path) between 0 m and 2 m which corresponds to the cycle path width.

Regarding the arrival time, the grid choice needs to be generic and comparable between all cyclists. Since each cyclist enters with a different speed, and the arrival time is a function of the cycling speed, a normalised relative temporal grid should be used that covers speed heterogeneity. One normalised indicator is the time difference inflicted compared to maintaining the speed they cycle at when the decision is made. This is hereafter referred to as “delay”. Positive delays indicate a later arriving time, which means that the cyclist is braking, while negative delays lead to an earlier arrival and, thus, accelerating behaviour.

In order to set up the temporal grid, the boundaries should be determined as well as the time step between the alternatives. In Figure 5.5 the observed delay compared to the initial cycling speed is visualised in relation to the lateral position upon arrival. Three types of symbols are used to separate three cyclist groups.

The black dots correspond to major stream cyclists and are mostly concentrated around delays of 0 s. So, as expected, these cyclists continue cycling with their original speed, following a zero-acceleration strategy, without being affected by other (crossing) cyclists.

The red crosses correspond to minor stream cyclists whose delays range from -4 s to +24 s. As some of the cyclists decide to yield by stopping, the duration of their stop is visible in this observed delay. The stopping duration, however, goes beyond the scope of the yielding decision model. For this reason, the delays observed by minor stream cyclists who do not stop are additionally noted with a blue circle. Their delay values go up to +6 s. Therefore, in the choice grid the delays are capped at +7 s, an upper threshold that stands for all delays higher than 7 s.

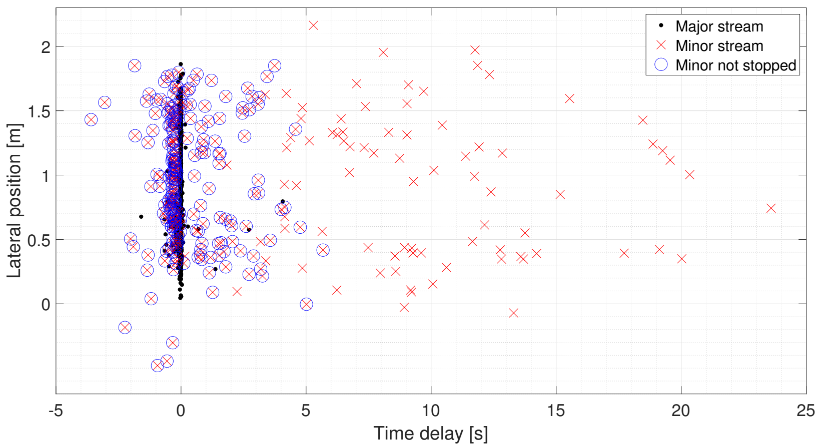


Figure 5.5: Observed combinations of delay to arrive at the crossing and lateral position upon arrival. The cycle path ranges between 0 and 2 m, while the 0 s delay corresponds to maintaining the speed the cyclist cycles at when making the yielding decision. Positive delays correspond to deceleration decisions and negative delays to acceleration.

Regarding the lower limit of the grid, it is set at 3 s. That is because the majority of the negative delays (acceleration decisions) goes up to 2 s, so an extra second is allowed for the few values corresponding to even greater acceleration. The time step is set at 0.5 s, because that is the average clearance time of the crossing for the major flow.

Based on these analyses, the choice set is constructed. It consists of combinations of a sublane (left/right) and a delay ranging from -3 s to +7 s with a step of 0.5 s. The size of the choice set of the operational mental layer model is, thus, 42 alternatives.

Following the definition of the choice set, the attributes that determine the attractiveness (utility) of an alternative need to be specified.

We hypothesise that the choice for a particular delay depends on the general attitude of cyclists towards acceleration (X_{Acc}) and deceleration (X_{Dec}), but also on the magnitude of the resulting delay separately considered for positive and negative delays X_{Tdec} and X_{Tacc}, respectively.

Regarding the choice for a sublane, we expect, based on the least effort assumption, that cyclists will want to keep their initial sublane. This means, for example, that if they make the decision on the left sublane, they prefer to arrive at the crossing on the left sublane (X_{LinLout}).

Apart from these attributes, the decision of minor stream cyclists is expected to be influenced by interactions occurring with cyclists in front of them heading in the same direction and cyclists coming from the major stream. An interaction with cyclists in the front takes place when there are stopped cyclists in one sublane. The fact that other cyclists have yielded by stopping might make it more attractive for the decision maker to also stop (and adhere to social norms) or might make it attractive to change lane to overtake the stopped cyclist (X_{BlockChL}). The conflict with cyclists coming from the right depends on the anticipation of that flow and the time that the crossing area will be occupied. In this application, perfect prediction of the occupancy is assumed, which means that the times when each major stream cyclist is present in the crossing area are known. These times are derived directly from the trajectory data of the major stream cyclists. We define two attributes to capture the anticipated conflict if the considered delay, and thus arrival time, would be selected. When only one major stream cyclist is expected to be in the crossing area, that is captured in the dummy attribute $X_{\mathrm{conflictS}}$. If more major stream cyclists are expected to be in the crossing area, their total number is reflected in the attribute $X_{\mathrm{conflictM}}$.

The reason why the presence of a single cyclist is differentiated from that of more cyclists is the fact that the minor stream cyclist (decision maker) is able to pass directly behind the major stream cyclist, while the latter is still in the crossing area. In other words, the conflict in this situation is easily resolved within the crossing area.

Along with the properties of the alternatives, we hypothesise that the traffic situation that the cyclists encounter play a role in their decision. These situations are represented by the class assigned to each cyclist as discussed in subsection 5.3.3. Therefore, interaction terms are created between the aforementioned properties of the choice alternatives and the cyclist classes.

Last but not least, the model should capture the attractiveness of the stopping alternatives (maximum delay value of 7 s) for those cyclists who have a conflict when following a zero-acceleration strategy and for whom the maximum delay would result in such a deceleration that they have to come to a full stop. The full stop is determined by assuming a minimum speed threshold beyond which balancing on the bicycle is not possible without using the feet. This threshold is set at 3.4 km/h. This threshold was determined such that the vast majority of stopped cyclists, based on the ground truth data, are also labelled as stopping. An alternative specific constant is thus added for those cyclists to the 7 s delay alternatives (X_{Stop}).

Figure 5.6 demonstrates the choice alternatives for which the different alternative specific constants are active. In the green cells (negative delays) X_{Acc} is active. In the orange cells X_{Dec} is active as well as in the 7 s delay alternatives. The latter also has X_{Stop} in its utility and is represented by a dark red colour in the figure. The 0 s delay alternatives have no alternative specific constant assigned to them.

These attributes describe the properties of all alternatives in the choice set. The choice set should, however, be adjusted to represent what is feasible for each cyclist. This feasibility is modelled via availability conditions. Alternatives, for example, that result in a simultaneous arrival (in time and space) at the cross line of the decision maker and preceding cyclists are made unavailable for the decision maker and succeeding cyclists. Moreover, some acceleration values result in an arrival time that would require travelling back in time compared to when the decision is made, or impossibly high speeds and are therefore considered not feasible and are excluded from the corresponding choice set.

Next to the availability conditions, for the model estimation we hypothesise that there is correlation between the stopping and the non-stopping alternatives provided in the choice set. For this reason, we create two nests: the "stop" nest consists of the 7 s delay for arrival time in either of the sublanes, and the "go" nest that comprises all the rest.

Furthermore, we assume that there is no serial correlation in the yielding decisions made by the same person in each lap. The justification behind this assumption is that the yielding choice is not of habitual nature. Instead we believe that it depends on the traffic situation that the cyclist encounters in each lap at the crossing, so intrapersonal heterogeneity prevents habitual behaviour. The consequence of this assumption is that the behavioural bias of each person in, for example, taking risks or giving priority, is ignored and several observations of a certain behaviour might lead to the conclusion that many cyclists exhibit it, while in reality it represents a smaller group of cyclists. As the unavailability of person identification in the dataset prevents us from assigning different decisions to the same cyclist and testing this assumption, future research should look into this effect in more detail.

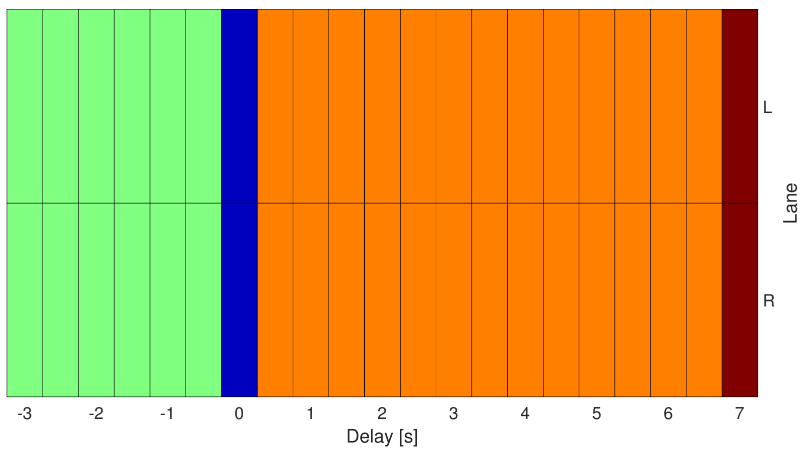


Figure 5.6: Assignment of alternative specific constants to choice alternatives: X_{Acc} in green cells, X_{Dec} in orange cells and 7 s delay alternatives, X_{Stop} in dark red cells.

Table 5.1: List of variables for the operational mental layer.

Attribute	**Unit**	**Explanation**
Z_{p}	-	dummy for major stream cyclists
Z_{n}	-	dummy for minor stream cyclists who do not have a conflict when following a zero-acceleration strategy
Z_{r}	-	dummy for minor stream cyclists who have a conflict when following a zero-acceleration strategy
X_{Acc}	-	dummy for negative delays
X_{Dec}	-	dummy for positive delays
X_{Tacc}	s	magnitude of negative delay
X_{Tdec}	s	magnitude of positive delay
X_{RinRout}	-	dummy for keeping the right sublane
X_{LinLout}	-	dummy for keeping left sublane
X_{BlockChL}	-	dummy for presence of stopped cyclists in the other sublane
$X_{\mathrm{ConflictS}}$	-	dummy for one conflicting cyclist in the crossing area at the considered arrival time
$X_{\mathrm{ConflictM}}$	cyclists	number of cyclists (at least two) in the crossing area at the considered arrival time
X_{Stop}	-	dummy for maximum delay alternatives (7 s in both sublanes)

5.4.2 Operational physical layer estimation approach

Within this layer, only the non-stopping cyclists are considered. That is because, as previously explained, the cyclists who decide to stop need to make a queuing position choice in the mental layer before their trajectory can be generated by the physical layer. For the non-stopping cyclists, the mental layer provides as input to the physical layer the intended lateral position and time to arrive at the cross line. In the physical layer, the cyclist decides in every time step the changes in pedalling and steering until the next time step. Through the sequence of time steps, and corresponding decisions, the cyclist trajectory is constructed.

The observed choices at every decision moment (time step) of all non-stopping cyclists are visualised by the black dots in Figure 5.7. The angular sections represent changes in steering angle per time step, while the arched zones correspond to changes in speed. The figure shows that in all observations the speed changes range between [-4,4] km/h, while most changes in steering angle are between [-15,15] degrees. Larger changes in the steering angle mostly occur in combination with a small change of speed. The threshold for the steering angle in both directions is set to 45 degrees, as only seven data points are beyond this angle and half of them are very close to 45 degrees. As the majority of the observations is within [-15,15] degrees, the discritesation of the steering angle is coarser for angles outside of this range and finer for angles within it.

The grid is, thus, designed to give the options for speed changes of -4 km/h to +4 km/h with a step of 1 km/h, and for steering angle changes of {-45,-30,-15,-10,-5,0,5,10,15,30,45} degrees. The combinations of these values of the grid comprise the choice set which has a size of 99 alternatives.

The attributes that are hypothesised to influence the choice relate to the willingness of cyclists to make specific changes in speed and angle within a time step. In the mental layer, their overall attitude towards speeding up or down is estimated for their entire trajectory upstream of the crossing. In the physical layer, this attitude becomes more specific by choosing whether to (i) make one intense manoeuvre or several smaller ones to achieve the intended final outcome; and (ii) combine the changes of speed and angle in a time step or apply the changes separately in different time steps. For this reason we define four attributes that separately examine the attitude towards accelerating (X_{pedal}), braking(X_{brake}), steering left (X_{steerL}) and steering right (X_{steerR}).

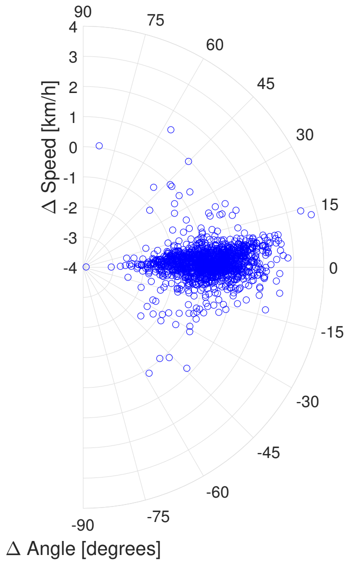


Figure 5.7: Observed changes in steering angle and speed per time step for all non-stopping cyclists.

Moreover, we expect that cyclists want to make changes that bring them closer to their final desired position at the desired time. So three attributes are defined to capture the remaining lateral deviation to the intended position (X_{Ydiff}) and the difference in arrival time if the new speed is adopted. As the effect of an early arrival (X_{early}) might be that the crossing area is still occupied, its contribution to the decision might differ from a late arrival (X_{late}) at an empty crossing. This effect is reversed in case of a cyclist who decides to speed up and accept a gap before the crossing area gets occupied.

Another attribute that is expected to affect the choices is whether the alternative leads the cyclist outside of the cycle path (X_{offpath}). Since cyclists were instructed to cycle within the marked lane as much as possible, this attribute should reveal the situations when it was preferable to leave the path than to follow the instructions.

Regarding the presence of other cyclists, in the physical layer only cyclists in front of the decision maker play a role. That is because the front cyclists act as obstacles or lead the way, thereby affecting the path to be followed. Based on the behavioural findings of Gavriilidou et al. (2019a), the behaviour against static obstacles differs from that towards moving persons. For this reason, in this application different attributes are created to test this effect once more. Moreover, the distance to cyclist closest to the decision maker had been found significant, and the longitudinal and lateral dimensions were separately evaluated. This separation is adopted in this model as well. Last but not least, the maximum difference in cycling speed for moving cyclists is considered as an attribute (X_{dVmov}), since the greater the difference, the more the situation resembles an interaction with a stationary obstacle.

In this layer the cyclist classification is again introduced in the form of interaction terms between the aforementioned attributes and the class of the decision maker. The underlying hypotheses are that major stream cyclists are (i) less willing to deviate from their desired arrival time as they have priority and (ii) less reactive to the movements of other major stream cyclists as they expect them to also proceed unhindered. The difference in behaviour of the minor stream cyclist groups is expected to be in the reaction to other cyclists, as when there is a conflict there might be more stopping or decelerating cyclists and so the minor stream cyclists with conflict should cycle more cautiously. The differences in the steering behaviour between the three groups is expected to reveal whether the curve has any remaining effects for minor stream cyclists (comparison of major stream to minor stream without conflict) and whether the presence of a conflict leads those cyclists to more steering manoeuvres (comparison between two minor stream groups).

In terms of availability conditions, we assume that cyclists do not move backwards, so negative speeds are not allowed. This means that depending on the actual speed of the cyclist at each decision moment, alternatives that would result in a negative cycling speed are made unavailable.

For the model estimation, it is assumed that there is no serial correlation in the decisions made by the same person at different time steps. This creates the risk of inconsistent behaviour between time steps (e.g. first accelerating and then braking). However, such behaviour could occur also in a real cycling situation and there is no reason to assume habitual behaviour or bias towards a particular choice. Therefore, a multinomial logit model is estimated for the physical layer.

Table 5.2: List of variables for the operational physical layer.

Attribute	Unit	Explanation
Z_{p}	-	dummy for major stream cyclists
Z_{n}	-	dummy for minor stream cyclists who do not have a conflict when following a zero-acceleration strategy
Z_{r}	-	dummy for minor stream cyclists who have a conflict when following a zero-acceleration strategy
X_{pedal}	m/s^2	magnitude of acceleration when considering positive speed changes
X_{brake}	m/s^2	magnitude of acceleration when considering negative speed changes
X_{steerL}	rad/s	magnitude of steering intensity when considering changes towards the left
X_{steerR}	rad/s	magnitude of steering intensity when considering changes towards the right
X_{Ydiff}	m	remaining lateral deviation from the intended arrival position
X_{early}	s	magnitude of positive difference between the intended arrival time and the arrival time given the considered speed change
X_{late}	s	magnitude of negative difference between the intended arrival time and the arrival time given the considered speed change
X_{offpath}	-	dummy for riding outside of the cycle path in the next step
X_{dXmov}	m	minimum longitudinal distance to moving cyclists in front
X_{dYmov}	m	minimum lateral distance to moving cyclists in front
X_{dVmov}	m/s	maximum speed difference between considered speed and speed of moving cyclists in front
X_{dXstop}	m	minimum longitudinal distance to stopped cyclists in front
X_{dYstop}	m	minimum lateral distance to stopped cyclists in front

5.5 Results and discussion

This section provides and discusses the model estimation results for each layer. Models have been estimated using Python Biogeme (Bierlaire 2016). A simulation is then performed with both estimated models to validate the corresponding behavioural level.

5.5.1 Operational mental layer model

A nested logit model was estimated using the nests described in section 5.4.1. The resulting nest coefficient was, however, not statistically significantly different in the two nests, so the hypothesis that these alternatives are nested in the choice set is rejected.

Following this finding, a multinomial logit model is estimated. The estimated coefficient values of the best performing model are shown in Table 5.3, for each of the cyclist classes. Cells that are empty had insignificant values. The final model consists of 15 parameters, which are all significant at 95% confidence level. The estimated parameter values are further discussed and interpreted for each cyclist class.

For major stream cyclists there is a strong negative attitude towards both acceleration ($\mathrm{ASC_{acc}} = -2.95$) and deceleration ($\mathrm{ASC_{dec}} = -5.60$), which means they prefer to keep their speed and confirms the assumption that cyclists are effort minimisers. Moreover, given the magnitude of the coefficient values, the major stream cyclists are more likely to accelerate than to decelerate. The other two attributes that are significant for this stream relate also to attitude, in this case regarding lane keeping. These cyclists have a preference to arrive at the crossing in the same sublane as the one they have when making a decision. The coefficient for the staying in the right sublane is slightly more positive than that of staying in the left sublane ($\beta_{\mathrm{RinRout}} = 7.03$ and $\beta_{\mathrm{LinLout}} = 6.46$, respectively). The difference of these values is statistically significant, which shows that cyclists comply to the general traffic rule of keeping to the right.

The behaviour of minor stream cyclists is different, depending on whether they have a conflict or not, which confirms the hypothesis related to the cyclist classes. Minor stream cyclists without a conflict have a positive attitude towards acceleration ($\mathrm{ASC_{acc}} = 2.23$) that reduces as the magnitude of the (negative) delay increases ($\beta_{\mathrm{Tacc}} = 1.88$).

Table 5.3: Estimated yielding model parameters and corresponding robust standard error for the operational mental layer. The three last columns differentiate the interaction terms of the three cyclist classes.

Coefficient name	**Coefficient value (Robust standard error)**		
	Major stream	Minor stream no conflict	Minor stream with conflict
$\mathrm{ASC}_{\mathrm{acc}}$	-2.95 (0.26)	2.23 (0.35)	
β_{Tacc}		1.88 (0.40)	
$\mathrm{ASC}_{\mathrm{dec}}$	-5.60 (0.20)	-0.94 (0.35)	1.51 (0.35)
β_{Tdec}		-0.27 (0.11)	-0.26 (0.12)
β_{LinLout}	6.46 (0.23)	0.54 (0.19)	
β_{RinRout}	7.03 (0.59)	1.65 (0.53)	
$\beta_{\mathrm{ConflictM}}$			-5.15 (0.08)
$\beta_{\mathrm{BlockChL}}$			-10.30 (0.21)
$\mathrm{ASC}_{\mathrm{stop}}$			2.51 (0.27)

All coefficient values are significant at 95% confidence level.

So the higher the acceleration, the less attractive it becomes. The delay beyond which acceleration is not any more attractive is -1.2 s (found by solving the equation $\mathrm{ASC}_{\mathrm{acc}} + \beta_{\mathrm{acc}} \cdot X_{\mathrm{Tacc}} = 0$). Additionally, they have a negative attitude towards deceleration, which gets stronger as the amount of delay increases.

These cyclists, similar to the major stream, have a tendency to keep their sublane that is stronger for the right than for the left sublane, but much weaker compared to the major stream cyclists. The reason we observe this weaker effect is that half of the minor stream cyclists switch lane and arrive on the right sublane. The vast majority of the minor stream is on the subleft lane in the beginning of the straight stretch due to the curve that is upstream. As they are unhindered, they take the inner side of the curve which leads them on the left sublane at the moment the mental moment is activated. Half of them stick to their entry lane, while the other half swerves to the right.

Cyclists in the minor stream that would encounter a conflict if they kept their cycling speed at the moment the decision is made, have a positive attitude towards deceleration that decreases along with the amount of incurred delay. The break-even point in this case is at $\mathrm{ASC}_{\mathrm{dec}}/\beta_{\mathrm{Tdec}} = 5.8$ s. Longer delays lead to a negative utility.

Along with these attitudinal variables, interaction with other cyclists influence their choices. When there is one or more cyclists stopped in the other sublane, they have a very strong deterrent from changing into that sublane ($\beta_{\mathrm{BlockChL}} = -10.3$).

The presence of more than one major stream cyclist at the crossing is another strong deterrent for minor stream cyclists, and gets even less attractive as the number of major stream cyclists expected in the crossing area increases ($\beta_{\mathrm{ConflictM}} = -5.15$). In case there is only one major stream cyclist, the coefficient of the $X_{\mathrm{ConflictS}}$ attribute was found to be insignificant. This is reasonable, since when interacting with only one major stream cyclist it is still possible (and safe) to arrive at the cross line and pass right behind this cyclist without experiencing a conflict.

Last but not least, the alternative specific constant for the maximum delay of 7 s has a positive coefficient ($\mathrm{ASC}_{\mathrm{stop}} = 2.51$) for those cyclists that have a conflict when following the zero-acceleration strategy and choosing this delay forces them to come to a full stop.

5.5.2 Operational physical layer model

The estimated values of the coefficients of the best performing model for the operational physical layer are provided in Table 5.4, for each of the cyclist classes. Cells that are empty had insignificant values. The final model consists of 23 parameters, which are all significant at 95% confidence level. The estimated parameter values are further discussed and interpreted, while comparing the three cyclist classes, starting with the those that express attitude towards speed and steering angle changes, followed by attraction to the intended arrival time and lateral position. Then, the parameters that describe the importance placed to interactions with the infrastructure and with other cyclists, are presented.

In this model, the attitudinal variables are consistent among all cyclist classes in terms of sign, only the magnitude differs. There is a negative association with all types of changes (towards higher and lower speeds, and towards steering to the left and to the right). It should be noted that the coefficients for braking (β_{brake}) and steering to the right (β_{steerR}) have positive values but the corresponding attributes have negative values, thus combined they produce a disutility when cyclists have to steer to the right or brake. Regarding the magnitude, minor stream cyclists who face a conflict are less deterred from braking, which is reasonable as they are expected to react and give way to the major stream. This means that avoiding the conflict is more important than having to decelerate.

Table 5.4: Estimated yielding model parameters and corresponding robust standard error for the operational physical layer. The three last columns differentiate the interaction terms of the three cyclist classes.

Coefficient name	**Coefficient value (Robust standard error)**		
	Major stream	Minor stream no conflict	Minor stream with conflict
β_{pedal}	-2.21 (0.16)	-1.07 (0.07)	-2.26 (0.18)
β_{brake}	4.72 (0.62)	3.71 (0.20)	1.55 (0.13)
β_{steerL}	-10.60 (0.97)	-2.57 (0.21)	-2.43 (0.25)
β_{steerR}	9.99 (1.07)	5.49 (0.49)	4.19 (0.49)
β_{Ydiff}		-6.67 (0.33)	-4.16 (0.50)
β_{early}		-2.79 (0.29)	-1.11 (0.14)
β_{late}			-0.39 (0.09)
β_{offpath}	-2.94 (0.61)	-1.52 (0.68)	
β_{dXmov}		-4.86 (0.83)	-1.42 (0.66)
β_{dXstop}		6.17 (1.26)	
β_{dVmov}		-3.34 (0.48)	

All coefficient values are significant at 95% confidence level.

Steering actions in both directions are equally unattractive for major stream cyclists, while steering to the right is less favourable for minor stream cyclists. This discrepancy in the two streams is attributed to the presence of the curve upstream of the straight stretch for the minor stream and the desire of cyclists to keep to the inner circle of the curve. As we found in the mental layer, there is a tendency to swerve to the right sublane when entering on the left. This is not by means of steering to the right, but rather by keeping the initial angle after exiting the curve. This angle is kept until the cyclist reaches the right sublane and then steers left to stabilise the position within the right sublane.

In addition to that, minor stream cyclists do not like to deviate from their intended lateral position ($\beta_{\text{Ydiff}} < 0$), nor their intended arrival time. Early arrivals are less attractive than late ones ($\beta_{\text{early}} < \beta_{\text{late}}$). An explanation for this behaviour is that an early arrival is linked to acceleration, which is unattractive in itself, but also the crossing area might be still occupied. This means that cyclists pick the first available passing opportunity, to inflict the least amount of delay on themselves, and that arriving a bit later than intended is still safe. These attributes were removed from the utility function of major stream cyclists due to high (more than 70%) correlation with the attitudinal variables towards any changes. These correlations led to unrealistic coefficient values that could not be independently interpreted.

Another finding is that cyclists, particularly in the major stream, avoid alternatives that lead them outside of the cycle path ($\beta_{\text{offpath}} < 0$). This is in line with their general attitude of least effort and minimal changes as there is no benefit to be gained, such as travel time savings, by straying off path. For minor stream cyclists without a conflict this attribute is less strong, but still inflicting a disutility. In the third cycling class this attribute is not a factor considered in the choices, which is reasonable since minor stream cyclists with a conflict also encounter stopped cyclists in their stream and might need to go out of the path to pass them.

Similar to the mental layer, in the physical layer minor stream cyclists take into account in their decision process the presence of other cyclists. It appears that only distances in the longitudinal direction have an effect. Those cyclists that do not have a conflict in front of them are strongly attracted to moving cyclists ($\beta_{\text{dXmov}} = -4.86$) and are repelled from stopped cyclists ($\beta_{\text{dXstop}} = 6.17$). The former shows a follow-the-leader effect, while the latter ensures no collision with stopped cyclists. Furthermore, minor stream cyclists without a conflict avoid increasing their speed compared to other cyclists, which is reasonable as they are following the leader and increasing their speed could result in unsafe situations, such as rear collisions.

Cyclists who have a conflict and do not stop, pay attention only to moving cyclists, while also displaying a follow-the-leader effect ($\beta_{\mathrm{dXmov}} = -1.42$).

5.5.3 Face validation using simulation

Using the estimated parameters for each model, a simulation is performed, where a prediction is made for each observation in the dataset. In this case, we have used the same dataset for estimation and face validation, as we do not have another dataset available and the dataset is too small to segment. Having the same dataset means that the attributes and availability conditions describing the situation at which every individual made a decision are the same as in the observed data. The simulation computes all utility functions and the corresponding probabilities of each alternative per individual. These probabilities are aggregated by averaging over all individuals to whom the corresponding alternative was available. The true (observed) choices are then visually compared with the predicted (simulated) probabilities for a specific choice.

The comparison for the operational mental layer is visualised in Figure 5.8. The observed choices are shown on the left and the simulated ones on the right of the figure. Each subplot corresponds to a different cyclist class. A different scale is used per class to best represent the variability of the observations over the alternatives. On top is the major stream, whose behaviour is well captured by the model. The vast majority of major stream cyclists opts, as expected, for the zero delay alternative with a higher preference for the right sublane.

Minor stream cyclists without a conflict (middle plots in Figure 5.8) have a preference for a small acceleration followed by a preference for zero delay. The developed model captures well this higher utility for small acceleration, and in some cases it allows for slightly larger accelerations as well. The light blue colour in the observed probabilities for the 7 s delay alternatives is attributed to misclassification of some stopped cyclists. For those, the speed was probably very low already, such that no conflict would take place if they maintained that speed, but actually they would have to accelerate to cycle comfortably or decelerate further to come to a full stop.

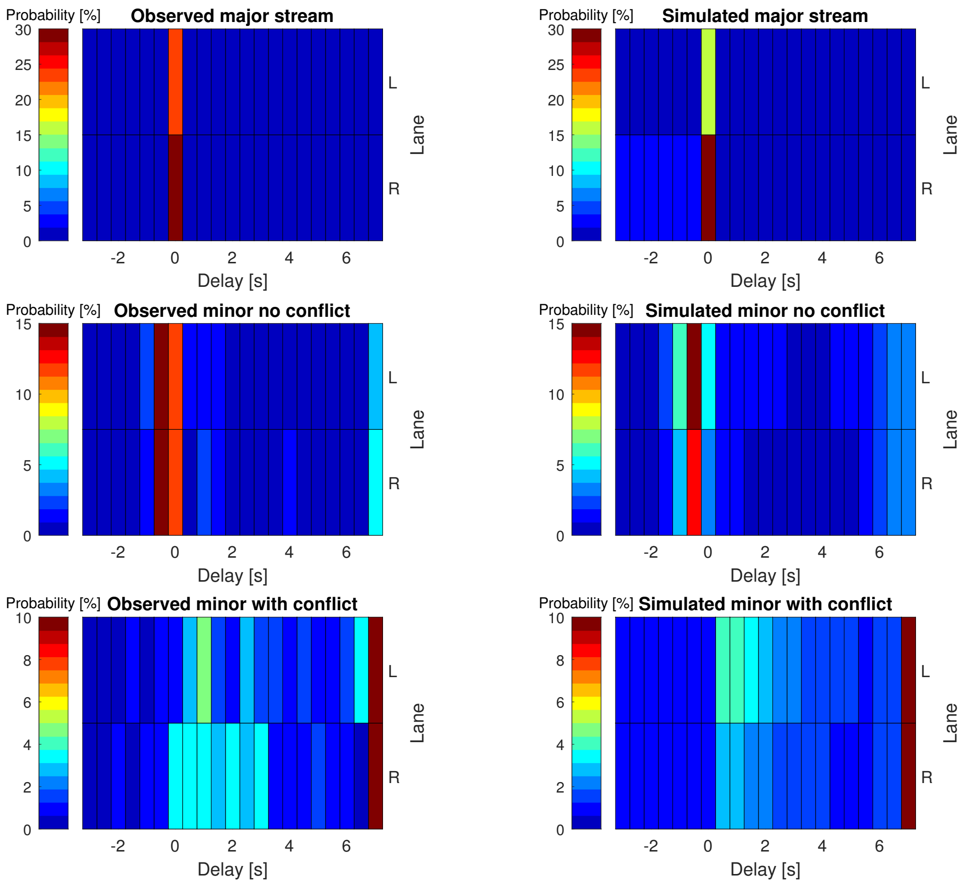


Figure 5.8: Probability of an alternative being selected in the observed (left) and the simulated (right) dataset separated per cyclist class. Note that a different scale is used for each class.

The bottom plots in Figure 5.8 correspond to the probabilities of minor stream cyclists with a conflict. In the observations, the majority of them yields by coming to a full stop, while the others have a preference for the right lane and decelerate with delays up to 3 s. The simulation captures the high share of stopping choices as well as the small delays (up to 2 s) for the rest.

The simulation results show a good match with the observations and, therefore, demonstrate the face validity of the developed model. Further validation should be performed using a different dataset.

The comparison for the operational physical layer is visualised in Figure 5.9, where the observations are in the top and the simulation results at the bottom. Overall, there is a limited range within the choice set that cyclists opt for regardless of their class, and the majority of the choices corresponds to either changes in speed or in steering angle, while keeping the other one constant.

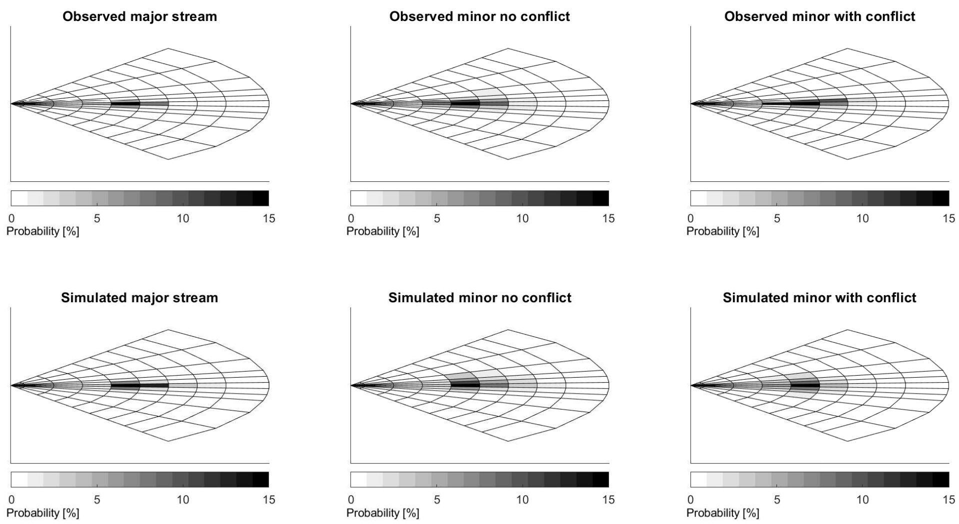


Figure 5.9: Probability of a combination of change in steering angle and speed to be selected in the observed (top) and the simulated (bottom) dataset separated per cyclist class.

On the left, the behaviour of major stream cyclists is displayed. The most likely choice is, both in the observations and for the simulation, that of no change in speed or angle, followed by a preference for small acceleration without steering and then by small steering to the left or to the right while maintaining the same speed. Where the two results diverge is in the probability for higher accelerations. In the simulation, it is more likely to accelerate than in reality but the probability is very low.

The minor stream cyclists without conflict are shown in the middle column. The pattern is almost identical between observations and simulation, with the majority of cyclists choosing for no change of state and the second highest probability going for a small steering action to the left. The main discrepancy is observed in the probability for small deceleration which is higher in the observations than in the simulation.

The simulation of minor stream cyclists with a conflict captures the movement to the left and decelerating actions. However, the probability assigned to these choices is smaller in the simulation than in the observations, which means they are selected less often in the simulation.

In order to further quantify the effect these deviations on the cyclist trajectories, the absolute percentage error made in each observation i is separately calculated for the longitudinal ($x_{\mathrm{error_i}}$) and the lateral ($y_{\mathrm{error_i}}$) direction. This is based on the simulated values ($x_{\mathrm{sim_i}}$, $y_{\mathrm{sim_i}}$) and the observed ones ($x_{\mathrm{obs_i}}$, $y_{\mathrm{obs_i}}$). The following formulas are used:

$$\begin{aligned} x_{\mathrm{error_i}} &= \frac{|x_{\mathrm{sim_i}} - x_{\mathrm{obs_i}}|}{x_{\mathrm{obs_i}}} \\ y_{\mathrm{error_i}} &= \frac{|y_{\mathrm{sim_i}} - y_{\mathrm{obs_i}}|}{y_{\mathrm{obs_i}}} \end{aligned} \tag{5.1}$$

The mean absolute percentage error is then calculated over all the observations within each cyclist class. The results are summarised in Table 5.5. It can be seen that the major stream has the smallest overall error in both directions, which is expected as they make the least amount of changes in their speed and steering angle. Minor stream cyclists with conflict have the highest error. This was also expected by the inability of the model to capture the higher probability for steering to the left and decelerations. However, all values are considered low and prove the face validity of the estimated model.

5.6 Conclusions and recommendations

In this chapter, cycling behaviour at bicycle crossings has been investigated and an operational model was estimated and face validated. The model development followed the two-layer framework proposed by Gavriilidou et al. (2019a) after improving it and substantiated its generalisability to interactions upstream unsignalised bicycle crossings. The

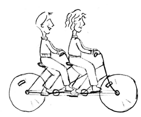


Table 5.5: Mean absolute percentage error in the simulated trajectories compared to the observed ones for each cyclist class.

Mean absolute percentage error	Cyclist classification		
	Major stream	Minor stream	
		No conflict	Conflict
x_{error}	1.84%	3.79%	4.18%
y_{error}	1.04%	2.52%	2.72%

improvement entailed an addition to the framework that allows a sequence of choices within the mental layer, prior to decisions in the physical layer. In the situation upstream of a crossing, cyclists first select their desired arrival time and lateral position at the crossing (mental layer). If the chosen time implies a stop, then a queuing position choice, also in the mental layer, is required. Otherwise, the trajectory is directly determined in the physical layer.

Using trajectory data from a controlled cycling experiment two discrete choice models were estimated, one for each layer. In the model development it was hypothesised that the behaviour depends on the traffic situation. Based on that cyclists were divided into three classes, namely major stream, minor stream without conflict and minor stream with conflict. Whether cyclists have a conflict or not was defined based on a zero-acceleration strategy by the decision maker and the anticipation of the occupancy of the crossing area. According to this definition, there is a conflict when a minor stream cyclist would arrive at the crossing area at a moment that it is occupied by a major stream cyclist, while having kept the cycling speed at the moment the decision to arrive at that time was made. The estimated models confirmed that the behaviour differs between these classes and revealed the influencing attributes.

In both layers it is observed that major stream cyclists want to make the least changes possible, both in terms of speed and angle. As they have priority in the crossing, it is reasonable that they are not willing to accommodate the crossing of minor stream cyclists and they just continue with their initial speed in their initial sublane.

Minor stream cyclists tend to accelerate when they have no conflict, but they decelerate at the presence of a conflict, even coming to a complete stop. Cyclists who decide to stop to avoid the conflict need to then decide their stopping position, so they go through another mental layer process that has previously been modelled (Gavriilidou et al. 2019a). For those cyclists that have a conflict and do not stop, the attributes in the physical layer revealed a follow-the-leader behaviour. Cyclists maintain a close distance to their predecessor and are careful not to increase the speed difference as that could lead to rear collisions.

Based on these behavioural insights from the estimated models, it is concluded that cyclists behave in a safe way and adhere to traffic rules by yielding to cyclists coming from the right. Also, they are efficient in their interactions as they select their arrival moment just as the major stream cyclist exits the crossing area. As the present dataset does not contain personal information, the model has not accounted for the effect of observing several decisions made by the same person (i.e. panel data). The model should then be extended before generalising the finding that cyclists behave in a safe and efficient way in pure bicycle interactions, and the validity of the developed models should be tested on more datasets. If this conclusion is generalised, then the design implication for cycling infrastructure is that bicycle crossings should be unsignalised. Moreover, the use of markings to indicate priority is unnecessary as long as priority rules are clear, and therefore intersecting cycle paths should not have additional priority marking or signs.

Some of the results pertaining to the operational physical layer are, however, already validated, as they were observed at the queue formation process upstream of a traffic light (Gavriilidou et al. 2019a). Cyclists, regardless of the situation, seem to be consistent in deterring from leaving the cycle path and in behaving differently towards stopped and moving cyclists. In order to further generalise the conclusions drawn for the cyclist population of the controlled experiment, its representativeness can be checked in future research by comparing the queue formation process upstream of the unsignalised crossing in the controlled experiment to the real-world observations at the signalised intersection.

Future research directions also pertain to the investigation of the validity of the model under different bicycle densities (e.g. more cyclists coming from the minor stream), configurations of the crossing (e.g. at a T-junction or intersections under an angle) and in more complex situations where multiple streams are present at the crossing (e.g. bidirectional traffic or turning cyclists).

Last but not least, the interaction between the two decision layers could be strengthened by allowing the re-evaluation of the decision made in the mental layer. In the present model, a single decision is made in the mental layer as the cyclist approaches the crossing under the assumption of perfect anticipation skills. However, the inaccuracies in the prediction of future positions of other cyclists in reality might call for correcting actions or an update of the initial yielding decision. Further investigation is thus required into the updating frequency or triggers, as well as the anticipation capabilities of the cyclists.

Chapter 6

Conclusion

The research in this dissertation focused on generating knowledge on microscopic operational cycling behaviour on dedicated cycling infrastructure, through data and models. The aim was to develop a mathematical model that describes this behaviour.

To this end, a novel modelling framework was proposed and its application was demonstrated by developing mathematical models that describe the decision making process of individual cyclists upstream of signalised and unsignalised intersections. Cyclist trajectory data were used to calibrate and validate these models. As such data were scarce prior to this research, an extensive data collection through a bicycle experiment was performed and followed by data processing to derive trajectories and data analyses to generate behavioural insights. The empirical findings from these data analyses and the calibrated parameters of the developed operational behaviour models are brought together in this final chapter to draw conclusions and make design recommendations for dedicated cycling infrastructure. Moreover, implications for practice are discussed and directions for future research are provided.

6.1 Main findings and conclusions

The main findings of this dissertation are summarised by answering the research questions that were formulated in the introduction in order to meet the research objective.

1. *Which modelling framework captures cycling decisions and movements?*

Decisions and movements that cyclists make while cycling and interacting with other cyclists and with the infrastructure are viewed as two intertwined processes, a mental and a physical, within the operational cycling behavioural level. The mental process corresponds to path choices that cyclists make within a route, such as whether to stop at a traffic light and if so, where to queue, whether to yield to cyclists with priority or accept a gap to cross an intersection. Based on the path choice, the physical process then determines the bicycle control dynamics through changes made by the cyclist by pedalling and steering.

A two-layer modelling framework captures these processes of the operational cycling behaviour. In this framework, a layer is dedicated to the decisions made in each process. At the same time, since the processes are intertwined, the layers are linked and exchange information relevant to the decisions to be made in the other layer.

This dissertation demonstrated the plausibility of the framework through an application and its generalisability through another application. The generalised framework shows that it is possible to have a sequence of different mental processes prior to the physical ones.

2. *Which are the key factors affecting the different decisions made while cycling?*

One of the reasons for opting for a two-layer framework was that decisions in the two layers are affected by different factors. Furthermore, decisions in the mental layer depend on the type of interaction and the traffic situation. This means that each decision in the mental layer has its own influencing factors. Decisions in the physical layer, on the other hand, are governed by the least effort principle.

The key factors affecting decisions in the mental layer have been found by means of a literature study, a survey and data analyses. The literature study revealed that the gap acceptance of cyclists is influenced by the speed of the cyclist and of the interacting traffic participant, as well as by the gap size. Regarding the choice to stop at a red light, both previous literature on red light running and the conducted survey showed that it depends on the traffic conditions (such as amount of conflicting traffic and number of queuing cyclists), but also on weather conditions and personal characteristics (such as nationality and familiarity with the crossing). Overtaking decisions, according to the survey, are additionally affected by time pressure, the slope and directionality of the cycle path and the bicycle type. Cyclists facing a yielding decision take into account several factors pertaining to the traffic conditions and their personal characteristics. Out of the list obtained from the survey, the data analyses confirmed the effect of the traffic conditions, and in particular of the speed, number and direction of approaching cyclists. Last but not least, the factors affecting the queuing position decision were found through data analyses to be the distances to the stop line, to the sidewalk and to the nearest bicycle in front.

Regarding the physical layer, the steering and pedalling decisions are influenced by the edges of the cycle path, as cyclists avoid leaving it. Additionally, they are attracted by the intended position selected in the mental layer. Last but not least, it is important whether there are stopped or moving cyclists in the vicinity, and if there are cyclists nearby, the speed difference and distance to them.

3. *Which datasets are needed to obtain the influence of the key factors on cycling decisions?*

Even though a survey can determine whether a specific factor is influencing a particular mental decision or not, the extent of the influence can only be determined by observations of individual cycling behaviour. For decisions within the physical layer, such observations are the only means of determining the influencing factors. Trajectories, i.e. sequence in time of individuals' positions in space, are the data type necessary to study this individual cycling behaviour on the operational level. More specifically, it is required to have trajectory data with high temporal resolution (less than a second) and spatial accuracy (a few centimetres) to capture the movements and interactions of individual cyclists with each other and with the infrastructure.

4. *To what extent do the key factors influence the different decisions?*

Discrete choice models have been estimated and validated that describe cycling behaviour upstream of signalised and unsignalised intersections. These models capture the decision making processes both in the mental and the physical layer, and reveal the extent to which the different factors influence the corresponding decisions. The specific weight (importance) of the influence is reflected in the coefficient value of the corresponding attribute and can be found in the model results presented in sections 2.6 and 5.5. Here, rather than listing all the factors, the interpretation of these weights is given, along with the behavioural insights gained from the estimated queuing and yielding models. To answer the question for other types of decisions, further research is required to unravel the extent of the influence of key factors on those decisions.

When deciding on a queuing position upstream of a red traffic light, the greatest influencing factor for the first arriving cyclist is the position of the 'request-green' button, since being able to press it requires stopping next to it. For the subsequent cyclists, the presence of a resting position (such as the curb of the sidewalk) is an attractive factor. As a result, cyclists prefer to stop on the right lane rather than build the queue in all lanes equally. An exception to this are the front spots on the left lane, because they are close to the stop line. Regarding the distance to other bicycles in the queue, cyclists leave the least space possible behind each other. This shows that cyclists value their comfort (resting position) more than their travel time, and their travel time more than their personal space.

The conclusion on the importance of comfort is confirmed also by the yielding behaviour model upstream of an unsignalised intersection. Cyclists in the major stream are opting for the least effort strategy which is to maintain their lane and cycling speed while approaching the crossing area. Minor stream cyclists behave similarly in the absence of a conflict, but when a cyclist that has priority is present, the minor stream cyclists yield. Yielding takes place by decelerating, and when necessary, by coming to a complete stop.

In the physical layer, the decision made in the mental layer is leading. This decision is an intended time moment and/or position to reach an intermediate destination. A trajectory is, then, created towards it, through changes in the cycling speed and steering angle at regular time intervals. Other influencing factors are the cyclists present on the cycle path and the properties of the infrastructure. In terms of interactions with other cyclists, the behaviour has been found to differ depending on whether the decision maker intends to stop or not and whether the others are stopped or moving. Those cyclists that intend to stop prefer to stay closer to other stopped cyclists, while those that keep moving avoid stopped cyclists and follow those that are moving ahead of them. In both cases attention is paid to the resulting speed differences, which are kept low. This indicates a safe behaviour, as rear collisions are avoided. Regarding the interaction with the infrastructure, cyclists try to stay within the edges of the cycle path. An exception to this is when there is a conflict with another cyclist and they do not want to stop, but instead they deviate from the path. Strong deterrents from getting off the cycle path are the change of surface type and especially the presence of a height difference, such as that of the curb of the sidewalk.

5. *Which design implications stem from the empirical findings and behavioural insights of the derived datasets and models?*

The behavioural insights obtained from the derived models confirm the assumption that cyclists are effort minimisers while abiding by the traffic rules. Cyclists stay within the designated infrastructure, so the first design recommendation is to ensure that the edges of the cycle path are clearly indicated. Further means to ensure that cyclists will not deviate from the infrastructure allocated to them are the change of type or height of the adjacent surface. Just upstream of signalised intersections an elevated right angle curb is advised, so that cyclists can use it as a resting spot. The elevation there is even recommended on both sides of the cycle path to achieve dense queues and, consequently, shorter discharge times.

Another empirical finding is that cyclists respect priority rules and behave in a safe way in bicycle-to-bicycle interactions. This has been investigated in T-junctions and four-legged crossings with one-way bicycle traffic. Though intersections have previously been found to be critical locations in a network for bicycle safety, the findings of the research within this dissertation show that bicycle intersections are not safety critical. In addition to this, the findings have implications for bicycle intersection design. More specifically, it is recommended to leave bicycle intersections unsignalised, as well as unequipped with traffic signs and priority markings. As long as the traffic rules are clear with respect to priority, cyclists adhere to them and additional markings or signs are unnecessary.

Regarding the design of the T-junction, the empirical findings suggest that a turn with a radius of 2 m would accommodate all merging cyclists. This is smaller than the minimum radius advised by CROW (2016), which is 5 m so that there is no need for a speed reduction. However, the safe interactions that have been observed in the T-junction are, at least partially, attributed to the speed reduction upstream of the conflict area, which make the approach of cyclists more cautious. Therefore, a radius that forces cyclists to decelerate without stopping is preferred.

Last but not least, it has been found that cyclists achieve smoother and more efficient interactions in a self-organised manner rather than when guided to a certain behaviour by the infrastructure design. The intervention that was specifically investigated was the use of lane marking upstream of the T-junction to guide minor stream cyclists to the left lane and give sufficient space for merging cyclists to fit in the right lane. This, however, limited the available lane width at all times, while in the absence of the lane marking cyclists made dynamic use of the width of the cycle path; cyclists used the full width when no encounter was about to take place and switched to the left lane when a cyclist with priority was approaching and both could continue uninterrupted.

6.2 Implications for practice

The research in this dissertation has been based on three pillars, namely models, data and design recommendations. The contributions within each of these have implications for practice.

In terms of data one contribution is the rich dataset of cyclist trajectories that has been collected. The corresponding practical implication is the possibility to further exploit this dataset to understand and model cycling behaviour in situations different from those elaborated on in this dissertation, as well as macroscopic characteristics of bicycle flows. Additionally, a guide has been provided on the process that should be followed to set-up and perform a (large-scale) data collection on operational behaviour. In this particular application the data collection was in the form of a controlled experiment, but the techniques and principles are also suitable and applicable for real-world observations. Moreover, they can be applied not only for bicycle traffic but for other transport modes as well. The most important lesson learnt is that a detailed plan is required, along with flexibility in adjusting it during the data collection, as unexpected events will occur regardless of the thorough planning, and behavioural assumptions made while planning will be found not to hold.

Next to the data, the developed operational cycling behavioural models constitute an important milestone for practice. Based on these mathematical models, it becomes possible to develop a complete bicycle traffic microsimulation model, covering scenarios on straight stretches, as well as at signalised and unsignalised intersections. Particularly at intersections, the simulation model can assess different designs of the cycle path both in terms of capacity and safety. The developed queuing model sheds light into the bicycle queue formation process at signalised intersections. Applications within the microsimulation entail the assessment of the design of bicycle lanes upstream of traffic signals, as well as of the stopping area. Depending on the cyclist demand, the shape and size of the stopping area might need to be adjusted, such that the spill-back of the bicycle queue does not cause further hindrance. Elements may also be added in the design to encourage denser queues and shorter discharge times. Another measure that can be evaluated is the duration of the green times for each stream, in relation to the expected demand and the available space. The yielding model captures cyclist interactions at unsignalised intersections with a focus on their yielding behaviour. Using the simulation, the capacity of the intersection can be checked in relation to the cycling demand, to know the point at which the queue of cyclists who stop to yield becomes too long and requires intervention. Moreover, different surrogate safety measures can be calculated to assess the safety of the interactions when their number increases. Another application of the microsimulation is the calculation of the interaction duration between cyclists in close proximity. That could be useful in epidemiological studies, assessing the spreading of a virus, such as COVID-19, the adequacy of the infrastructure and the need for additional measures.

Regarding the design recommendations for dedicated cycling infrastructure, several conclusions can be drawn from the aforementioned microsimulations. The ones that stem directly from the research within this dissertation have been listed in the answer to the last research question provided in the previous section. The key points are that (i) no markings or signs are necessary to indicate priority or guide cyclists in a specific position on cycle path, and (ii) elevated right angle curbs are advised just upstream of signalised intersections to serve as a resting spot while queuing and make the queues denser.

6.3 Directions for future research

Despite all its merits and contributions, the research performed within this dissertation is simply a stepping stone towards developing an understanding of operational cycling behaviour and models that capture and replicate this behaviour.

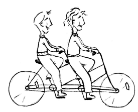
Further research is required to advance the models that have been estimated. More specifically, advancements can be sought in the decision moments and exchange rate of information between the mental and the physical layers. In the applications so far, a sequential feed-forward approach was followed. This means that a single decision was made in the mental layer and executed in the physical layer without reconsiderations. The moment when this first decision is made in the mental layer should be further investigated. Moreover, as the traffic situation changes and the anticipation skills of the cyclist at that single moment are not perfect, it is reasonable to allow for feedback loops, either event or frequency-based, which may lead to a revision of the decision made in the mental layer.

The models should also be extended to take cyclist heterogeneity into account, as it has been shown to affect the efficient use of the infrastructure. Heterogeneity could be included either in the form of additional attributes or using more advanced logit models. Examples of additional attributes are personal characteristics, such as cycling experience and age, the bicycle type and the trip purpose, all of which are expected to have a direct effect on the desired cycling speed. In the present research, multinomial logit models were estimated, which assume the same degree of influence of the different attributes among different persons and also ignore any serial correlation potentially present in consecutive choices of the same cyclist. They could be enhanced by allowing the coefficients to follow a distribution rather than have a single value. Moreover, if the decisions made by the same person can be identified and labelled in the data, the models can be extended to account for the panel data.

Another direction for future research is the development of the models currently defined in the two-layer framework. For example, the queuing position choice upstream of a traffic light should be preceded by a choice to stop, or not. The latter, also called red light running, is a confirmed phenomenon and should, thus, not be considered negligible and ignored by the models. The yielding model at a bicycle crossing has been estimated in a situation where only two streams are present and all cyclists continue their paths straight. It would be valuable to investigate whether the findings pertaining to safe interactions and respect for priority rules hold in more complex situations, where more streams are present, cyclists are turning and there is bi-directional traffic.

Apart from these extensions to existing models, and in order to complete the modelling suite, models should be estimated to capture overtaking manoeuvres and the path that cyclists follow when turning, either in a curve or when merging with another bicycle stream.

Last but not least, the rich dataset that has been collected has its limitations and future research could try to compensate for those. One drawback is the artificial environment and controlled conditions, a remedy for which would be to make observations in the real-world of similar traffic situations. Another limitation is that all data has been collected considering a 2 m wide cycle path, which is not always available. A comparison considering the effect of the cycle path width but also of its surface type, on cycling behaviour is therefore advised. Finally, it would be interesting to compare the cycling behaviour observed in different countries, as the cycling culture and good infrastructure found in the Netherlands might explain the findings of respectful and safe behaviour better than the role of being a cyclist in traffic.



Appendix A

Data smoothing

In this appendix, the results of smoothing with different sliding window lengths are presented, as well as of the homogenisation with different time steps. The original, smoothed and final data points are displayed in the plots of Figure A.1 for the cyclist trajectories, Figure A.2 for the cycling speed and Figure A.3 for the steering angle.

Sliding window lengths, denoted by k, of 3 and 6 are compared, which means that every data point is replaced by the mean value of its 3 or 6 surrounding frames. The values are selected based on the average frame rate of 6fps, which means that frames of 0.5 or 1 second are used for the smoothing. The only difference that can be detected in the plots is in the region of -9 m in the x direction where the two cameras overlap. In the original dataset there seems to be a jump from the back camera to the front, which is smoothed with this process. The results with $k = 6$ convert this jump to an almost continuous trajectory and are therefore favourable.

Regarding the time step, denoted by dt, values of 0.5 and 1 second are compared, as they are considered to be reasonable time intervals for a new decision to be made. Smaller values would coincide with the frame rate, while larger ones would lead to very few points per trajectory as the average trajectory duration is 7 seconds. The difference in this comparison can be observed in the plots for the speed and the angle where the larger dt is shown to be better at muting the noise in the dataset which is introduced due to the manual tracking. For this reason, the time step is chosen at 1 second.

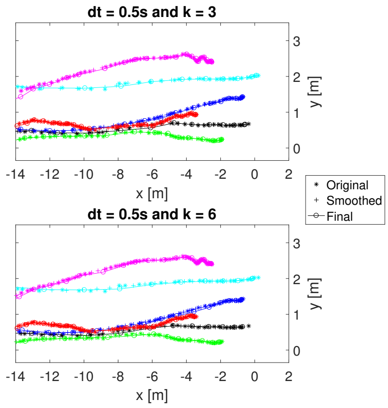


Figure A.1: Cyclist trajectories during a red light phase in the original dataset, when smoothed with different sliding window lengths (top: $k = 3$ and bottom: $k = 6$) and finally when homogenised with different time steps (left: $dt = 0.5s$ and right: $dt = 1s$). The point (0,0) is the location where the stop line meets the curb of the sidewalk.

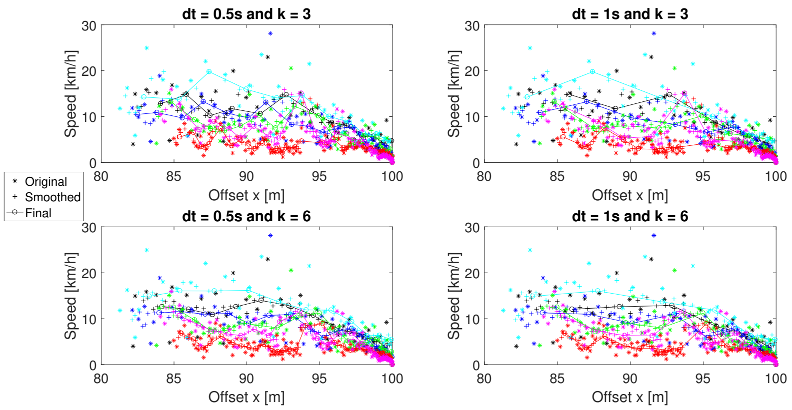


Figure A.2: Cyclist speed when approaching a red traffic light in the original dataset, when smoothed with different sliding window lengths (top: $k = 3$ and bottom: $k = 6$) and finally when homogenised with different time steps (left: $dt = 0.5s$ and right: $dt = 1s$). The positions in x have been offset such that they end at the same point for all cyclists.

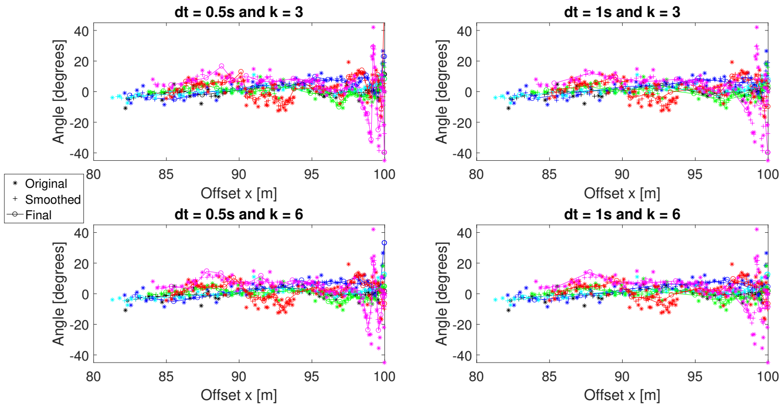


Figure A.3: Cyclist steering angle when approaching a red traffic light in the original dataset, when smoothed with different sliding window lengths (top: $k = 3$ and bottom: $k = 6$) and finally when homogenised with different time steps (left: $dt = 0.5s$ and right: $dt = 1s$). The positions in x have been offset such that they end at the same point for all cyclists.

Bibliography

Andresen, E., M. Chraibi, A. Seyfried, F. Huber (2013) Basic driving dynamics of cyclists, in: *Simulation of Urban MObility User Conference*, Springer, pp. 18–32.

Antonini, G., M. Bierlaire, M. Weber (2006) Discrete choice models of pedestrian walking behavior, *Transportation Research Part B: Methodological*, 40(8), pp. 667–687.

Avineri, E., F. Steven (2013) Has the introduction of the cycle-to-work scheme increased levels of cycling to work in the uk?, in: *Transportation Research Board 92nd Annual Meeting*, pp. 13–20.

Baum, L. (2008) Smart trips summit-u: an individualized marketing approach to changing travel behavior, *TDM Review*, (4).

Bierlaire, M. (2016) Pythonbiogeme: A short introduction, *Ecole Polytechnique Federale, School of Architecture, Civil and Environmental Engineering. Lausanne, Switzerland: Transport and Mobility Laboratory.*

Botma, H., H. Papendrecht (1991) Traffic Operation Of Bicycle Transporting, *Transportation Research Record: Journal of the Transportation Research Board*, 1320, pp. 65–72.

Chen, X., B. Lin, H. Han (2013) Characteristics of mixed non-motorized traffic flow: A comparative analysis with motorized and pedestrian traffic flow, *Transportation Research Board 92nd Annual Meeting Compendium of Papers.*

Choudhury, C. F., M. Ben-Akiva, M. Abou-Zeid (2010) Dynamic latent plan models, *Journal of Choice Modelling*, 3(2), pp. 50–70.

Costa, M., A. Bichicchi, M. Nese, C. Lantieri, V. Vignali, A. Simone (2019) T-junction priority scheme and road user's yielding behavior, *Transportation Research Part F: Traffic Psychology and Behaviour*, 60, pp. 770–782.

CROW (2016) *Ontwerpwijzer fietsverkeer*, CROW.

Daamen, W., S. P. Hoogendoorn (2003) Experimental research of pedestrian walking behavior, *Transportation Research Record: Journal of the Transportation Research Board*, 1828(1), pp. 20–30.

Dodge, Y. (2008) *The concise encyclopedia of statistics*, Springer Science & Business Media.

Dozza, M., J. Werneke (2014) Introducing naturalistic cycling data: What factors influence bicyclists' safety in the real world?, *Transportation Research Part F: Traffic Psychology and Behaviour*, 24, pp. 83 – 91.

Dubuy, V., K. De Cocker, I. De Bourdeaudhuij, L. Maes, J. Seghers, J. Lefevre, K. De Martelaer, G. Cardon (2013) Evaluation of a workplace intervention to promote commuter cycling: A re-aim analysis: Veerle dubuy, *The European Journal of Public Health*, 23(suppl_1), pp. ckt123–069.

Duives, D. C. (2016) *Analysis and modelling of pedestrian movement dynamics at large-scale events*, Ph.D. thesis, Delft University of Technology.

Field, A. (2013) *Discovering statistics using IBM SPSS statistics*, Sage.

Fishman, E., C. Cherry (2016) E-bikes in the mainstream: reviewing a decade of research, *Transport Reviews*, 36(1), pp. 72–91.

Flower, J., J. Parkin (2019) Understanding attitudes to priorities at side road junctions, *Transportation Research Part F: Traffic Psychology and Behaviour*, 62, pp. 246–257.

Fukuda, D., T. Seo, K. Yamada, H. Yaginuma, N. Matsuyama (2014) An econometric-based model of pedestrian walking behavior implicitly considering strategic or tactical decisions, in: *Pedestrian and Evacuation Dynamics 2012*, Springer, pp. 615–624.

Gavriilidou, A., W. Daamen, Y. Yuan, S. Hoogendoorn (2019a) Modelling cyclist queue formation using a two-layer framework for operational cycling behaviour, *Transportation Research Part C: Emerging Technologies*, 105, pp. 468–484.

Gavriilidou, A., W. Daamen, Y. Yuan, S. P. Hoogendoorn (2020a) To yield or not to yield? A behavioural model at unsignalised bicycle crossings, submitted to Transportation Research Part C: Emerging Technologies.

Gavriilidou, A., W. Daamen, Y. Yuan, N. van Nes, S. Hoogendoorn (2020b) Empirical findings on infrastructure efficiency at a bicycle T-junction, *Physica A: Statistical Mechanics and its Applications*, p. 125675.

Gavriilidou, A., M. J. Wierbos, W. Daamen, Y. Yuan, V. L. Knoop, S. P. Hoogendoorn (2019b) Large-scale bicycle flow experiment: setup and implementation, *Transportation Research Record: Journal of the Transportation Research Board*, 2673(5), pp. 709–719.

Gavriilidou, A., Y. Yuan, H. Farah, S. P. Hoogendoorn (2017) Microscopic cycling behavior model using differential game theory, in: *International Conference on Traffic and Granular Flow*, Springer, pp. 497–506.

Goñi Ros, B., Y. Yuan, W. Daamen, S. P. Hoogendoorn (2018) Empirical analysis of the macroscopic characteristics of bicycle flow during the queue discharge process at a signalized intersection, *Transportation Research Record: Journal of the Transportation Research Board*.

Greibe, P., T. S. Buch (2016) Capacity and Behaviour on One-way Cycle Tracks of Different Widths, *Transportation Research Procedia*, 15, pp. 122–136.

Heinen, E., B. Van Wee, K. Maat (2010) Commuting by bicycle: an overview of the literature, *Transport Reviews*, 30(1), pp. 59–96.

Helmers, G., L. Aberg (1978) Driver behavior in intersections as related to priority rules and road design, *An exploratory study.(Rapport 167). Linköping, Sweden: VTI.*

Hoogendoorn, S. P., P. H. Bovy (2002) Normative pedestrian behaviour theory and modelling, in: *Proceedings of the 15th international symposium on transportation and traffic theory*, pp. 219–245.

Hoogendoorn, S. P., P. H. L. Bovy (2004) Pedestrian route-choice and activity scheduling theory and models, *Transportation Research Part B: Methodological*, 38(2), pp. 169–190.

Hoogendoorn, S. P., W. Daamen (2005) Pedestrian behavior at bottlenecks, *Transportation Science*, 39(2), pp. 147–159.

Hoogendoorn, S. P., W. Daamen, P. H. Bovy (2003) Extracting microscopic pedestrian characteristics from video data, in: *Transportation Research Board Annual Meeting*, Citeseer, pp. 1–15.

Huang, L., J. Wu (2009) Cyclists' path planning behavioral model at unsignalized mixed traffic intersections in China, *IEEE Intelligent Transportation Systems Magazine*, 1(2), pp. 13–19.

Huang, L., J. Wu, F. You, Z. Lv, H. Song (2017) Cyclist social force model at unsignalized intersections with heterogeneous traffic, *IEEE Transactions on Industrial Informatics*, 13(2), pp. 782–792.

Hydén, C., Å. Svensson, L. Sakshaug (2007) Yielding behaviour and interaction at bicycle crossings, in: *TRB's 3rd Urban Street Symposium.*

Imbert, C., M. te Brömmelstroet (2014) The desire lines of bicycle users in amsterdam, *Copenhague: Copenhagenize Design Company.*

Jiang, H., T. Wen, P. Jiang, H. Han (2013) Research on cyclists microscopic behaviour models at signalized intersection, in: *16th International Conference Road Safety on Four Continents. Beijing, China (RS4C 2013). 15-17 May 2013.*

Jiang, R., M.-b. Hu, Q.-s. Wu, W.-g. Song (2016) Traffic Dynamics of Bicycle Flow: Experiment and Modeling, *Transportation Science*, 51(3), pp. 998–1008.

Jiang, R., M.-B. Hu, Q.-S. Wu, W.-G. Song (2017) Traffic dynamics of bicycle flow: experiment and modeling, *Transportation Science*, 51(3), pp. 998–1008.

Jin, S., M. Liu, L. Shen, D. Ma (2017) Modelling speed − flow relationships for bicycle traffic flow, *Proceedings of the Institution of Civil Engineers - Transport*, 170, pp. 194–204.

Knoppers, P., H. Van Lint, S. P. Hoogendoorn (2012) Automatic stabilization of aerial traffic images, in: *Transportation Research Board 91st Annual Meeting*.

Kucharski, R., A. Drabicki, T. Kulpa, A. Szarata (2017) Multichannel cyclist queuing behaviour at signalised cycle crossings, in: *Proceedings of 6th Symposium of the European Association for Research in Transportation*.

Kucharski, R., A. Drabicki, K. Żyłka, A. Szarata (2019) Multichannel queueing behaviour in urban bicycle traffic, *European Journal of Transport and Infrastructure Research*, 19(2).

Lee, T.-C., J. W. Polak, M. G. Bell (2009) New approach to modeling mixed traffic containing motorcycles in urban areas, *Transportation Research Record: Journal of the Transportation Research Board*, 2140(1), pp. 195–205.

Lee, T.-C., K. Wong (2016) An agent-based model for queue formation of powered two-wheelers in heterogeneous traffic, *Physica A: Statistical Mechanics and its Applications*, 461, pp. 199–216.

Li, M., F. Shi, D. Chen (2011) Analyze bicycle-car mixed flow by social force model for collision risk evaluation, in: *3rd International Conference on Road Safety and Simulation*, pp. 1–22.

Li, Z., M. Ye, Z. Li, M. Du (2015) Some Operational Features in Bicycle Traffic Flow, *Transportation Research Record: Journal of the Transportation Research Board*, 2520, pp. 18–24.

Liang, X., M. Baohua, X. Qi (2012) Psychological-physical force model for bicycle dynamics, *Journal of Transportation Systems Engineering and Information Technology*, 12(2), pp. 91–97.

Luo, Y., B. Jia, J. Liu, W. H. Lam, X. Li, Z. Gao (2015) Modeling the interactions between car and bicycle in heterogeneous traffic, *Journal of Advanced Transportation*, 49(1), pp. 29–47.

Ma, X., D. Luo (2016) Modeling cyclist acceleration process for bicycle traffic simulation using naturalistic data, *Transportation Research Part F: Traffic Psychology and Behaviour*, 40, pp. 130–144.

Mai, X., W. Lv, X. Wei, W. Song, R. Jiang (2013) Analyzing the characteristics of unidirectional bicycle movement around a track based on digital image processing, *Procedia Engineering*, 62, pp. 519–524.

Mallikarjuna, C., K. R. Rao (2009) Cellular automata model for heterogeneous traffic, *Journal of Advanced Transportation*, 43(3), pp. 321–345.

Mathew, T. V., O. Abdullah, C. R. Munigety, S. Anirudha (2012) Strip-based simulation model for mixed traffic conditions, Tech. rep.

Michon, J. A. (1985) A critical view of driver behavior models: What do we know, what should we do, *Human Behavior and Traffic Safety*, pp. 485–520.

Nabavi Niaki, M., N. van Nes, A. Gavriilidou, W. Daamen (2019) Safety effects of lane marking nudge at a bicycle intersection, Abstract presented at 8th Annual International Cycling Safety conference.

Navin, F. P. D. (1994) Bicyce Traffic Flow Characteristics: Experimental Results and Comparisons, *ITE Journal*, 64(3), pp. 31–37.

Ng, A., A. K. Debnath, K. C. Heesch (2017) Cyclist'safety perceptions of cycling infrastructure at un-signalised intersections: Cross-sectional survey of queensland cyclists, *Journal of Transport & Health*, 6, pp. 13–22.

Oketch, T. G. (2000) New modeling approach for mixed-traffic streams with nonmotorized vehicles, *Transportation Research Record: Journal of the Transportation Research Board*, 1705(1), pp. 61–69.

Olde Kalter, M. (2007) Vaker op de fiets? effecten van overheidsmaatregelen [more often the bicycle? effects of government measures].

Pai, C.-W., R.-C. Jou (2014) Cyclists' red-light running behaviours: An examination of risk-taking, opportunistic, and law-obeying behaviours, *Accident Analysis & Prevention*, 62, pp. 191–198.

Phillips, R. O., T. Bjørnskau, R. Hagman, F. Sagberg (2011) Reduction in car–bicycle conflict at a road–cycle path intersection: Evidence of road user adaptation?, *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(2), pp. 87–95.

Pucher, J., R. Buehler (2008) Making cycling irresistible: lessons from the Netherlands, Denmark and Germany, *Transport Reviews*, 28(4), pp. 495–528.

Pucher, J., R. Buehler (2012) City cycling (vol. 11).

Rasmussen, J. (1983) Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models, *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(3), pp. 257–266.

Richardson, M., B. Caulfield (2015) Investigating traffic light violations by cyclists in dublin city centre, *Accident Analysis & Prevention*, 84, pp. 65–73.

Schleinitz, K. (2016) *Cyclists' road safety-Do bicycle type, age and infrastructure characteristics matter?*, Ph.D. thesis, Technische Universität Chemnitz.

Shepherd, R. (1994) Road and path quality for cyclists, in: *Proc. 17th ARRB conference*, vol. 17, Gorld Coast, Queensland, pp. 133–147.

Shiomi, Y., T. Hanamori, M. Eng, U. Nobuhiro, H. Shimamoto (2012) Modeling traffic flow dominated by motorcycles based on discrete choice approach, in: *Proceedings of 1st LATSIS Conference*.

Svensson, Å., J. Pauna-Gren (2015) Safety at cycle crossings : The relationship between motor vehicle drivers' yielding behaviour and cyclists' traffic safety, in: *Proceedings of 28th ICTCT Workshop in Ashdod, Israel*.

Twaddle, H. (2017) *Development of tactical and operational behaviour models for bicyclists based on automated video data analysis*, Dissertation, Technische Universität München.

Twaddle, H., G. Grigoropoulos (2016) Modeling the speed, acceleration, and deceleration of bicyclists for microscopic traffic simulation, *Transportation Research Record: Journal of the Transportation Research Board*, 2587(1), pp. 8–16.

van Haperen, W., S. Daniels, T. De Ceunynck, N. Saunier, T. Brijs, G. Wets (2018) Yielding behavior and traffic conflicts at cyclist crossing facilities on channelized right-turn lanes, *Transportation Research Part F: Traffic Psychology and Behaviour*, 55, pp. 272–281.

Vansteenkiste, P., G. Cardon, E. D'Hondt, R. Philippaerts, M. Lenoir (2013) The visual control of bicycle steering: The effects of speed and path width, *Accident Analysis & Prevention*, 51, pp. 222–227.

Vasic, J., H. J. Ruskin (2012) Cellular automata simulation of traffic including cars and bicycles, *Physica A: Statistical Mechanics and its Applications*, 391(8), pp. 2720–2729.

Wang, D., D. Zhou, S. Jin, D. Ma (2015) Characteristics of mixed bicycle traffic flow on the conventional bicycle path, *Transportation Research Board 94nd Annual Meeting Compendium of Papers*.

Wexler, M. S., A. El-Geneidy (2017) Keep'em separated: Desire lines analysis of bidirectional cycle tracks in montreal, canada, *Transportation Research Record: Journal of the Transportation Research Board*, 2662(1), pp. 102–115.

Wijlhuizen, G., J. van Petegem, C. Goldenbeld, P. van Gent, J. de Bruin, J. Commandeur, V. Kars (2016) Development cyclerap instrument for safety assessment bicycle infrastructure: efficiency manual intensity measurements, reliability evaluation infrastructure, Tech. rep., SWOV.

Wu, C., L. Yao, K. Zhang (2012) The red-light running behavior of electric bike riders and cyclists at urban intersections in china: an observational study, *Accident Analysis & Prevention*, 49, pp. 186–192.

Yao, D., Y. Zhang, L. Li, Y. Su, S. Cheng, W. Xu (2009) Behavior modeling and simulation for conflicts in vehicles-bicycles mixed flow, *IEEE Intelligent Transportation Systems Magazine*, 1(2), pp. 25–30.

Yuan, Y., W. Daamen, B. Goñi-Ros, S. P. Hoogendoorn (2018) Investigating cyclist interaction behavior through a controlled laboratory experiment, *Journal of Transport and Land Use*, 11(1), pp. 833–847.

Zeegers, T. (2004) Width of bicycle paths, Detailed article by the fietsersbond about the desired width of different kinds of bicycle paths., Fietsersbond.

Zhang, J., W. Mehner, E. Andresen, S. Holl, M. Boltes, A. Schadschneider, A. Seyfried (2013) Comparative analysis of pedestrian, bicycle and car traffic moving in circuits, *Procedia – Social and Behavioral Sciences*, 104, pp. 1130–1138.

Zhang, R., J. Wu, L. Huang, F. You (2017) Study of bicycle movements in conflicts at mixed traffic unsignalized intersections, *IEEE Access*, 5, pp. 10108–10117.

Zhang, W., J.and Mehner, S. Holl, M. Boltes, E. Andresen, A. Schadschneider, A. Seyfried (2014) Universal flow-density relation of single-file bicycle, pedestrian and car motion, *Physics Letters, Section A: General, Atomic and Solid State Physics*, 378, pp. 3274–3277.

Zhang, Y., P. Stopher, B. Halling (2010) An evaluation of travelsmart tools for travel behaviour change, URL http://hdl.handle.net/2123/19283.

Zhao, Y., H. Zhang (2017) A unified follow-the-leader model for vehicle, bicycle and pedestrian traffic, *Transportation Research Part B: Methodological*, 105, pp. 315–327.

Zhou, D., C. Xu, D.-H. Wang, J. Sheng (2015) Estimating capacity of bicycle path on urban roads in hangzhou, china, *Transportation Research Board 94nd Annual Meeting Compendium of Papers*.

Summary

As the title suggests, cyclists are the main topic of this dissertation and more specifically, their behaviour while they are 'in motion'. The term 'in motion' is used in the title to represent microscopic operational cycling behaviour, which is the behaviour of cyclists, treated as individuals (microscopic level), while they are riding their bicycle and making decisions on how to interact with other traffic participants and with the infrastructure (operational level). Within this dissertation, models are developed to capture this behaviour using data collected for this purpose. Further empirical data analyses led to more behavioural insights and design recommendations were provided based on the findings. In this summary, each of these elements is shortly discussed, along with the need for this research.

To begin with, the motivation behind this topic is the growing interest in cycling in urban environments worldwide, and the accompanying need to provide appropriate infrastructure for cyclists. In order to determine which infrastructure is appropriate, such that it meets safety and comfort requirements, it is necessary to understand the behaviour and preferences of cyclists.

In general, the building blocks towards this understanding are data and models. Data in the form of empirical observations reveal preferences and typical characteristics of cycling behaviour, while mathematical models aim to capture and predict these. Then, with the development of simulation models on the basis of the mathematical ones, the performance of different infrastructure designs is assessed under varying traffic conditions, leading to recommendations regarding the most appropriate design.

Most research efforts so far have focused on mixed traffic situations, which means that cyclists are interacting with motorised vehicles, scooters or pedestrians. At the same time, it is found in literature, that in

an urban environment that envisages to promote bicycle use and safety, the provided infrastructure should be dedicated to cyclists. To this end, data and models are needed that capture bicycle-to-bicycle interactions on dedicated cycling infrastructure. Such data and models are the main contributions of this dissertation.

Within this dissertation, the first step towards acquiring data and models is the definition of the operational level and the behaviours that fall within. As the decisions and movements of cyclists while cycling and interacting with other cyclists and with the infrastructure are viewed as two intertwined processes, a mental (decisions) and a physical (movements), two layers are distinguished within the operational cycling behavioural level. In the operational mental layer, cyclists build up their paths within the route. Path choices refer, among other things, to yielding, accepting a gap to merge or cross, stopping for a red traffic signal, turning, and overtaking. The execution of each of these path choices is done in the operational physical layer, where bicycle control dynamics are applied by the cyclist in the form of changes in pedalling force and steering angle.

Having developed this two-layer modelling framework, the next step is the acquisition of cyclist movement data, namely trajectories. A trajectory is a sequence in time of an individual's positions in space. More specifically, in order to observe how cyclists interact and how they use infrastructure, trajectory data are needed with high temporal resolution (less than a second) and spatial accuracy (a few centimetres). In this dissertation, two trajectory datasets have been used: one from real-world observations previously collected from a high vantage point on a cycle path upstream of a signalised intersection in Amsterdam, and one from a controlled experiment collected within this research where several types of traffic situations were observed.

The controlled experiment has been set up to enable different types of interactions among the cyclists and ensure a sufficient amount of observations. The track was designed with long straight stretches to allow overtaking manoeuvres, but also with intersecting cycle paths to observe merging, crossing, stopping and queuing decisions. Overhead cameras were placed above the track to follow the positions of all cyclists on the track at any moment. Scenarios were then designed to capture different types of interactions and the influence of bicycle types. The schedule was constructed in such a way the variability in the type of interactions, the duration of the different runs and breaks in between runs kept the interest and stamina of the participants.

Following the data collection, analyses on the extracted trajectories are performed to obtain behavioural insights. In this dissertation, three situations at intersections are studied, namely the traffic efficiency at a bicycle T-junction, the yielding behaviour at a bicycle crossing and the queue formation process upstream of a red traffic light.

The traffic efficiency is evaluated using a framework that assesses the performance of individual cyclists, of the infrastructure use and of occurring interactions. The factors whose effect on the efficiency is investigated pertain to characteristics of the cyclists and their bicycles, but also to the addition of lane marking at the T-junction to guide the position of the cyclists and facilitate merging manoeuvres. The findings suggest that there is little to no effect on the efficiency following the introduction of the lane marking, while great effect results from the heterogeneity of the cycling population. It was observed that cyclists self-organise and make space for other cyclists to merge, also in the absence of guidance. The separation of the flow by lane markings, on the other hand, forced lane changes and resulted in longer queues and delays.

Regarding the yielding behaviour, an operational cycling model was estimated based on the data collected from the controlled experiment, following the developed two-layer framework. Cyclists who are arriving upstream of a crossing select their desired arrival time and lateral position at the crossing based on their anticipation of the cyclist movements in the conflicting direction. This takes place within the operational mental layer and is captured by a discrete choice model. The latter reveals that cyclists in the 'major' stream (i.e. with priority) follow a least effort strategy: they are not willing to accommodate the crossing of minor stream cyclists and they continue without changing lane or speed. Minor stream cyclists, on the other hand, adjust their behaviour: they accelerate when they have no conflict at the crossing, and they decelerate at the presence of a conflict, even coming to a complete stop if necessary.

For those cyclists that do not stop, a discrete choice model is estimated to capture their decisions in the physical layer. In this layer, several decisions are made in consecutive time steps that determine the trajectory of the cyclist towards the intended arrival time and position at the crossing. Interactions with cyclists who are also upstream of the crossing and in front of the decision maker are in this case important. The results show that cyclists maintain a close distance to their predecessor and are careful not to increase the speed difference as that could lead to rear collisions. This is evidence of a follow-the-leader behaviour.

For cyclists that come to a stop, discrete choice models are estimated to represent the queue formation process using the real-world observations upstream of a signalised intersection. In the mental layer, cyclists select their intended stopping position, and in the physical layer they determine their trajectory towards it in consecutive time steps. When the traffic light is red, the first arriving cyclist chooses to stop next to the 'request-green' button to be able to press it. In general, cyclists prefer to stop close to each other, but once the front stopping positions are occupied, there is a trend to stop closer to the curb of the sidewalk rather than build up all sublanes equally. This is intuitive because cyclists want to use the curb as a resting position when stopped and as an assist when starting to move again. At the same time, it shows that cyclists value their comfort (resting position) more than their travel time, and their travel time more than their personal space.

While cycling, the model results show that cyclists behave differently towards stopped and moving cyclists, which is reasonable since stopped cyclists form an obstacle on the way and an increase of the speed difference might lead to unsafe situations. Moreover, cyclists deter from changing surface type and, even more strongly, from cycling over and off curbs.

Based on the empirical findings from the data analyses and the behavioural insights from the estimated models, recommendations are made for the design of dedicated cycling infrastructure. The most important one is to leave bicycle intersections unsignalised, as well as unequipped with traffic signs and priority markings. As long as the traffic rules are clear with respect to priority, cyclists adhere to them and additional markings or signs are unnecessary. This is due to the self-organisation that has been observed, which leads to an efficient and flexible use of the infrastructure.

Despite all its merits and contributions, the research performed within this dissertation is a stepping stone towards developing an understanding of operational cycling behaviour and models that capture and replicate this behaviour. Further research is required to validate and extend the estimated models and to develop new ones for different types of interactions. Additionally, simulation models should be developed on the basis of these models to assess the performance of different infrastructure designs and ensure the provision of safe and comfortable cycling infrastructure. Last but not least, the representativeness of these findings in cyclist populations outside of the Netherlands should be tested, as the cycling culture and good infrastructure found in the Netherlands might explain the observed behaviour better than the role of being a cyclist in traffic.

Samenvatting

Dit proefschrift gaat over personen die deelnemen aan het verkeer als fietser. We onderzoeken hoe hun zogenoemde microscopische operationele fietsgedrag kan worden beschreven. Dit is het gedrag van fietsers, als individuen (microscopisch niveau), terwijl ze fietsen en beslissingen nemen hoe ze om moeten gaan met andere verkeersdeelnemers en met de infrastructuur (operationeel niveau). In dit proefschrift worden modellen ontwikkeld om dit gedrag te formaliseren met behulp van data die speciaal hiervoor zijn verzameld. Verdere empirische data analyses hebben geleid tot meer gedragsinzichten en op basis van de bevindingen zijn aanbevelingen voor het ontwerp van fietsinfrastructuur gedaan. In deze samenvatting wordt elk van deze elementen kort besproken, evenals de noodzaak van dit onderzoek.

De motivatie achter dit onderwerp is de groeiende belangstelling voor het fietsen in stedelijke omgevingen wereldwijd, en de daarmee gepaard gaande noodzaak om een passende infrastructuur voor fietsers aan te bieden. Om te bepalen welke infrastructuur voldoet aan de eisen van veiligheid en comfort van de fietser is het noodzakelijk om het gedrag en de voorkeuren van fietsers te begrijpen.

In het algemeen bestaan de bouwstenen voor dit inzicht uit data en modellen. Data in de vorm van empirische waarnemingen onthullen voorkeuren van fietsers en typische kenmerken van fietsgedrag, terwijl wiskundige modellen erop gericht zijn deze te formaliseren en te voorspellen. Op basis van de wiskundige modellen worden simulatiemodellen ontwikkeld, waarmee de prestaties van verschillende infrastructuurontwerpen worden beoordeeld onder verschillende verkeersomstandigheden. Dit leidt tot aanbevelingen voor het meest geschikte ontwerp.

De meeste onderzoeken hebben zich tot nu toe gericht op situaties met gemengd verkeer, hetgeen betekent dat fietsers interacteren met al dan niet gemotoriseerde voertuigen en/of voetgangers. Tegelijkertijd blijkt uit de literatuur dat in een stedelijke omgeving die het fietsgebruik en de veiligheid wil bevorderen, er een aparte infrastructuur moet worden gecreëerd voor fietsers. Hiervoor zijn data en modellen nodig die de interactie formaliseren tussen individuele fietsers op fietsinfrastructuur. Dergelijke data en modellen zijn de belangrijkste bijdragen van dit proefschrift.

De eerste stap richting het verzamelen van data en het ontwikkelen van modellen is het definiëren van het operationele niveau en het gedrag dat daarbinnen valt. De beslissingen en bewegingen van fietsers tijdens het fietsen en de interactie met andere fietsers en met de infrastructuur worden gezien als twee met elkaar verweven processen. In het operationele fietsgedrag worden dan ook twee lagen onderscheiden, namelijk een mentale laag waarbinnen de beslissingen vallen en een fysieke laag die de fietsbewegingen representeert. In de operationele mentale laag bepalen fietsers hun paden binnen de route. Beslissingen met betrekking tot dit pad hebben onder meer betrekking op het geven van voorrang, het accepteren van een hiaat om in te voegen of over te steken, het stoppen voor een rood verkeerslicht, het afslaan en het inhalen van een andere fietser. De uitvoering van elk van deze beslissingen vindt plaats in de operationele fysieke laag, waarbij de besturing van de fiets door de fietser wordt gerealiseerd in de vorm van veranderingen in de trapkracht en stuurhoek.

Na de ontwikkeling van dit tweelaagse modelkader is de volgende stap het verzamelen van data over fietsbewegingen, de zogenaamde trajectoriën. Een trajectorie is een sequentie in de tijd van de posities van een individu in de ruimte. Om de interactie tussen fietsers en hun gebruik van de infrastructuur te kunnen observeren, zijn trajectoriedata nodig met een hoge temporele resolutie (minder dan een seconde) en grote ruimtelijke nauwkeurigheid (enkele centimeters). In dit proefschrift zijn twee data sets met trajectoriën gebruikt: één data set met observaties die vanaf een hoog standpunt zijn verzameld op een fietspad stroomopwaarts van een kruispunt met verkeerslichten in Amsterdam, en één data set afkomstig uit een gecontroleerd experiment dat in het kader van dit onderzoek is uitgevoerd en waarbij verschillende verkeerssituaties zijn geobserveerd.

Het gecontroleerde experiment is opgezet om verschillende typen interacties tussen fietsers te kunnen observeren en te zorgen voor voldoende waarnemingen. Het parcours is ontworpen zowel met lange rechte stukken om inhaalmanoeuvres mogelijk te maken, als met kruisende fietspaden om beslissingen over invoegen, oversteken, stoppen en wachten te observeren. Boven het parcours zijn camera's opgehangen om de posities van alle fietsers op het parcours te kunnen volgen. Vervolgens zijn er scenario's ontworpen om verschillende typen interacties en de invloed van het type fiets vast te leggen. Het schema, met daarin de variabiliteit in het type interacties, de duur van de verschillende runs en de pauzes tussen de runs, is zo opgebouwd dat de interesse en het uithoudingsvermogen van de deelnemers behouden blijven.

Na het verzamelen van de data worden analyses op de trajectoriën van het gecontroleerde experiment uitgevoerd om inzichten in het fietsgedrag te verkrijgen. In dit proefschrift worden drie situaties op kruispunten bestudeerd, namelijk de efficiëntie van de verkeersstromen op een fiets-T-kruispunt, het gedrag om voorrang te geven op een fietskruispunt en het gedrag bij het vormen van een rij stroomopwaarts van een rood verkeerslicht.

De efficiëntie van de verkeersstromen op de fiets-T-kruispunt wordt geëvalueerd aan de hand van een raamwerk waarmee de prestaties van individuele fietsers, het gebruik van de infrastructuur en de optredende interacties worden beoordeeld. Voor verschillende factoren wordt het effect op de efficiëntie wordt onderzocht. Het gaat daarbij om factoren met betrekking op kenmerken van de fietsers en hun fiets, maar ook op de aanwezigheid van rijstrookmarkering op de T-splitsing om de fietsers te sturen en manoeuvres om in te voegen te vergemakkelijken. De bevindingen suggereren dat er weinig tot geen effect is op de efficiëntie na de invoering van de rijstrookmarkering, terwijl de heterogeniteit van de fietspopulatie een groot effect heeft. Geconstateerd is dat in fietsstromen, net als in voetgangersstromen, zelforganisatie optreedt, waarbij fietsers ruimte maken voor andere fietsers om in te voegen, ook zonder geleiding of markering. De scheiding van de stroom fietsers door middel van rijstrookmarkeringen dwong fietsers daarentegen tot het veranderen van rijstrook, hetgeen resulteerde in langere wachtrijen en vertragingen.

Aan de hand van het ontwikkelde raamwerk wordt een operationeel fietsmodel geschat voor het geven van voorrang op basis van de data die werden verzameld met het gecontroleerde experiment. Fietsers die stroomopwaarts van het kruispunt aankomen selecteren hun gewenste aankomsttijd en laterale positie op het kruisingsvlak op basis van hun anticipatie op de fietsbewegingen in de conflicterende richting. Dit gebeurt binnen de operationele mentale laag en wordt geformaliseerd in een discreet keuzemodel. De resultaten van het keuzemodel laten zien dat fietsers met voorrang een strategie van de 'minste effort' volgen: ze zijn niet bereid om het kruisen van fietsers zonder voorrang te accommoderen en ze fietsen door zonder van rijstrook te wisselen of hun snelheid te veranderen. De fietsers zonder voorrang daarentegen passen hun gedrag aan: ze versnellen als ze geen conflict hebben op het kruispunt, en ze vertragen als er een conflict is, en komen zo nodig zelfs helemaal tot stilstand.

Voor de fietsers die niet stoppen wordt een discreet keuzemodel geschat om hun beslissingen in de operationele fysieke laag vast te leggen. In deze laag worden in opeenvolgende tijdstappen verschillende beslissingen genomen die het pad van de fietser naar de beoogde aankomsttijd en positie op de kruising bepalen. Interacties met fietsers die ook stroomopwaarts van het kruispunt en vóór de fietser staan zijn in dit geval belangrijk. De resultaten laten zien dat fietsers dicht bij hun voorganger blijven en voorzichtig zijn met het vergroten van het snelheidsverschil, want dat kan leiden tot kop-staartbotsingen. Dit suggereert een 'follow-the-leader' gedrag.

Voor fietsers die tot stilstand komen worden discrete keuzemodellen geschat om de wachtrijvorming te formaliseren aan de hand van waarnemingen stroomopwaarts van een kruispunt met verkeerslichten. In de mentale laag selecteren fietsers hun beoogde stoppositie en in de fysieke laag bepalen ze in opeenvolgende tijdstappen hun pad daarheen. Als het verkeerslicht rood is, kiest de eerst aankomende fietser ervoor om naast de drukknop te stoppen, om deze in te kunnen drukken om groen licht aan te vragen. In het algemeen stoppen fietsers het liefst dicht bij elkaar, maar als de voorste stopplaatsen eenmaal bezet zijn, is er een trend om dichter bij de stoep te stoppen in plaats van de rijen in alle stroken gelijk op te bouwen. Hierdoor kunnen fietsers de stoep gebruiken als rustpositie bij het stoppen en als hulpmiddel bij het weer in beweging komen. Tegelijkertijd laat het zien dat fietsers meer waarde hechten aan hun comfort (rustpositie) dan aan hun reistijd, en meer aan hun reistijd dan aan hun persoonlijke ruimte.

Voor het fietsen zelf laten de modelresultaten zien dat fietsers zich verschillend gedragen ten opzichte van stilstaande en bewegende fietsers, wat redelijk is omdat stilstaande fietsers een obstakel vormen op de weg en een toename van het snelheidsverschil kan leiden tot onveilige situaties. Bovendien zijn fietsers niet geneigd om van type ondergrond te veranderen en, sterker nog, om over stoepranden te fietsen.

Op basis van de empirische bevindingen uit de data-analyses en de gedragsinzichten uit de geschatte modellen worden aanbevelingen gedaan voor het ontwerp van fietsinfrastructuur. Het belangrijkste is om fietskruispunten niet te voorzien van verkeerslichten, verkeersborden en voorrangsmarkeringen. Zolang de verkeersregels duidelijk zijn met betrekking tot de voorrang, houden de fietsers zich eraan en zijn extra markeringen of verkeersborden overbodig. Dit heeft te maken met de waargenomen zelforganisatie, die leidt tot een efficiënt en flexibel gebruik van de infrastructuur.

Het onderzoek in dit proefschrift is een opstap naar het ontwikkelen van inzicht in operationeel fietsgedrag en modellen die dit gedrag formaliseren en repliceren. Verder onderzoek is nodig om de geschatte modellen te valideren en uit te breiden en om nieuwe modellen te ontwikkelen voor verschillende typen interacties. Daarnaast moeten op basis van deze wiskundige modellen simulatiemodellen worden ontwikkeld om de prestaties van verschillende infrastructuurontwerpen te beoordelen en te zorgen voor een veilige en comfortabele fietsinfrastructuur. Ten slotte moet de representativiteit van deze bevindingen bij fietserspopulaties buiten Nederland worden getest, omdat de fietscultuur en goede infrastructuur die in Nederland worden aangetroffen het waargenomen gedrag beter zouden kunnen verklaren dan alleen het gebruik maken van de fiets in het verkeer.

About the author

Alexandra Gavriilidou was born in Thessaloniki, Greece, on 8 May 1990. In 2014 she obtained with distinction a Diplom-Ingenieur degree (equivalent to a joint Bachelor and Master degree) in Mechanical Engineering from the Aristotle University of Thessaloniki, Greece. For her graduation project she worked on optimising the power management of a Plug-in Hybrid Electric Vehicle (PHEV) at Toyota Motor Europe in Brussels, Belgium.

After her graduation, she pursued an MSc in Civil Engineering, track Transport and Planning, at the Delft University of Technology, which she obtained with distinction in 2016. Her master thesis was entitled "Transfer synchronisation of public transport services using passenger data".

In October 2016 she started her PhD research at the department of Transport and Planning, faculty of Civil Engineering at the Delft University of Technology. Her research was part of the Allegro project, a Horizon 2020 project funded by the ERC, aiming to unravel the behaviour of active modes using innovative data and models.

Since October 2020 she works at the department of Transport and Planning as lecturer and researcher. Her main research interests include operational cycling behaviour, data collection, behavioural analyses, and discrete choice modelling.

Publications

Journal papers

1. **Gavriilidou, A.** and Cats, O. (2019). Reconciling transfer synchronization and service regularity: real-time control strategies using passenger data. *Transportmetrica A: Transport Science*, 15(2), pp.215-243.

2. **Gavriilidou, A.**, Wierbos, M.J., Daamen, W., Yuan, Y., Knoop, V.L. and Hoogendoorn, S.P. (2019). Large-scale bicycle flow experiment: setup and implementation. *Transportation Research Record*, 2673(5), pp.709-719.

3. **Gavriilidou, A.**, Daamen, W., Yuan, Y. and Hoogendoorn, S.P. (2019). Modelling cyclist queue formation using a two-layer framework for operational cycling behaviour. *Transportation Research Part C: Emerging Technologies*, 105, pp.468-484.

4. **Gavriilidou, A.**, Daamen, W., Yuan, Y., van Nes, N. and Hoogendoorn, S.P. (2020). Empirical findings on infrastructure efficiency at a bicycle T-junction. *Physica A: Statistical Mechanics and its Applications*, p.125675.

5. **Gavriilidou, A.**, Daamen, W., Yuan, Y. and Hoogendoorn, S.P. (2020 under review). To yield or not to yield? A behavioural model at unsignalised bicycle crossings. Submitted to *Transportation Research Part C: Emerging Technologies.*

Conference contributions

1. **Gavriilidou, A.**, Cats, O., Leffler, D., Corman, F. and Hoogendoorn, S.P. (2017). Real-time Transfer Synchronization of Public Transport Services using Passenger Data. *In proceedings of the 96th Transportation Research Board Annual Meeting*, Washington DC.

2. **Gavriilidou, A.**, Yuan, Y., Farah, H. and Hoogendoorn, S.P. (2017). Microscopic Cycling Behavior Model Using Differential Game Theory. *In International Conference on Traffic and Granular Flow* (pp. 497-506). Springer, Cham.

3. **Gavriilidou, A.**, Yuan, Y., Farah, H. and Hoogendoorn, S.P. (2017). Face validation of a microscopic cycling behaviour model using differential game theory. Presented at the *Symposium of the European Association for Research in Transportation (hEART 2017)*, Haifa, Israel.

4. **Gavriilidou, A.**, Yuan, Y., Farah, H. and Hoogendoorn, S.P. (2018). Determinants of cyclists' operational decision making in the Netherlands. Presented at the *International Association for Travel Behaviour Research (IATBR 2018)*, Santa Barbara, California, USA.

5. **Gavriilidou, A.**, Wierbos, M.J., Daamen, W., Yuan, Y., Knoop, V.L. and Hoogendoorn, S.P. (2019). Large-scale bicycle flow experiment: set-up and implementation. *In proceedings of the 98th Transportation Research Board Annual Meeting*, Washington, D.C., USA.

6. **Gavriilidou, A.**, Susacasa, L.P., Reddy, N. and Daamen, W. (2020). Bicycle Parking Choice Behaviour at Train Stations. A Case Study in Delft, the Netherlands. *In Traffic and Granular Flow 2019* (pp. 479-485). Springer, Cham.

7. Knoop, V.L., Hänseler, F., Wierbos, M.J., **Gavriilidou, A.**, Daamen, W. and Hoogendoorn, S.P. (2020). Voronoi Densities for Bicylists: Adaptation for Finite Object Size and Speed. *In Traffic and Granular Flow 2019* (pp. 515-521). Springer, Cham.

8. **Gavriilidou, A.**, Daamen, W., Yuan, Y. and Hoogendoorn, S.P. (2019). Understanding cycling behaviour through large-scale experiments. Presented at the *Cycling Research Board (CRB 2019)*, Delft, The Netherlands.

9. Nabavi Niaki, M., van Nes, N., **Gavriilidou, A.** and Daamen W. (2019) Safety effects of lane marking nudge at a bicycle inter-

section, Presented at the *8th Annual International Cycling Safety conference (ISCS2019)*, Brisbane, Australia.

10. **Gavriilidou, A.**, Daamen, W., Yuan, Y., van Nes, N. and Hoogendoorn, S.P. (2020). Infrastructure efficiency of T-junctions for cyclists. *In proceedings of the 99th Transportation Research Board Annual Meeting*, Washington, D.C., USA.

11. Hoogendoorn, S.P., **Gavriilidou, A.**, Daamen, W. and Duives, D.C. (2021). Game theoretical framework for bicycle operations: a multi-strategy framework. To be presented at the *24th International Symposium on Transportation and Traffic Theory (ISTTT24)*, Beijing, China.

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series, For a complete overview of more than 250 titles see the TRAIL website: www.rsTRAIL.nl.

Gavriilidou, A., *Cyclists in Motion: From data collection to behavioural models*, T2021/7, February 2021, TRAIL Thesis Series, the Netherlands

Methorst, R., *Exploring the Pedestrians Realm: An overview of insights needed for developing a generative system approach to walkability*, T2021/6, February 2021, TRAIL Thesis Series, the Netherlands

Walker, F., *To Trust or Not to Trust? Assessment and calibration of driver trust in automated vehicles*, T2021/5, February 2021, TRAIL Thesis Series, the Netherlands

Schneider, F., *Spatial Activity-travel Patterns of Cyclists*, T2021/4, February 2021, TRAIL Thesis Series, the Netherlands

Madadi, B., *Design and Optimization of Road Networks for Automated Vehicles*, T2021/3, January 2021, TRAIL Thesis Series, the Netherlands

Krabbenborg, L.D.M., *Tradable Credits for Congestion Management: support/reject?*, T2021/2, January 2021, TRAIL Thesis Series, the Netherlands

Castelein, B., *Accommodating Cold Logistics Chains in Seaport Clusters: The development of the reefer container market and its implications for logistics and policy*, T2021/1, January 2021, TRAIL Thesis Series, the Netherlands

Polinder, G.J., *New Models and Applications for Railway Timetabling*,

T2020/18, December 2020, TRAIL Thesis Series, the Netherlands

Scharpff, J.C.D., *Collective Decision Making trough Self-regulation*, T2020/17, November 2020, TRAIL Thesis Series, the Netherlands

Guo, W. *Optimization of Synchromodal Matching Platforms under Uncertainties*, T2020/16, November 2020, TRAIL Thesis Series, the Netherlands

Narayan, J., *Design and Analysis of On-Demand Mobility Systems*, T2020/15, October 2020, TRAIL Thesis Series, the Netherlands

Gong, X., *Using Social Media to Characterise Crowds in City Events for Crowd Management*, T2020/14, September 2020, TRAIL Thesis Series, the Netherlands

Rijal, A., *Managing External Temporal Constraints in Manual Warehouses*, T2020/13, September 2020, TRAIL Thesis Series, the Netherlands

Alonso González, M.J., *Demand for Urban Pooled On-Demand Services: Attitudes, preferences and usage*, T2020/12, July 2020, TRAIL Thesis Series, the Netherlands

Alwosheel, A.S.A., *Trustworthy and Explainable Artificial Neural Networks for choice Behaviour Analysis*, T2020/11, July 2020, TRAIL Thesis Series, the Netherlands

Zeng, Q., *A New Composite Indicator of Company Performance Measurement from Economic and Environmental Perspectives for Motor Vehicle Manufacturers*, T2020/10, May 2020, TRAIL Thesis Series, the Netherlands

Mirzaei, M., *Advanced Storage and Retrieval Policies in Automated Warehouses*, T2020/9, April 2020, TRAIL Thesis Series, the Netherlands

Nordhoff, S., *User Acceptance of Automated Vehicles in Public Transport*, T2020/8, April 2020, TRAIL Thesis Series, the Netherlands

Winter, M.K.E., *Providing Public Transport by Self-Driving Vehicles: User preferences, fleet operation, and parking management*, T2020/7,

April 2020, TRAIL Thesis Series, the Netherlands

Mullakkal-Babu, F.A., *Modelling Safety Impacts of Automated Driving Systems in Multi-Lane Traffic*, T2020/6, March 2020, TRAIL Thesis Series, the Netherlands

Krishnakumari, P.K., *Multiscale Pattern Recognition of Transport Network Dynamics and its Applications: A bird's eye view on transport*, T2020/5, February 2020, TRAIL Thesis Series, the Netherlands

Wolbertus, *Evaluating Electric Vehicle Charging Infrastructure Policies*, T2020/4, February 2020, TRAIL Thesis Series, the Netherlands

Yap, M.D., *Measuring, Predicting and Controlling Disruption Impacts for Urban Public Transport*, T2020/3, February 2020, TRAIL Thesis Series, the Netherlands

Luo, D., *Data-driven Analysis and Modeling of Passenger Flows and Service Networks for Public Transport Systems*, T2020/2, February 2020, TRAIL Thesis Series, the Netherlands

Erp, P.B.C. van, *Relative Flow Data: New opportunities for traffic state estimation*, T2020/1, February 2020, TRAIL Thesis Series, the Netherlands

Zhu, Y., *Passenger-Oriented Timetable Rescheduling in Railway Disruption Management*, T2019/16, December 2019, TRAIL Thesis Series, the Netherlands

Chen, L., *Cooperative Multi-Vessel Systems for Waterborne Transport*, T2019/15, November 2019, TRAIL Thesis Series, the Netherlands

Kerkman, K.E., *Spatial Dependence in Travel Demand Models: Causes, implications, and solutions*, T2019/14, October 2019, TRAIL Thesis Series, the Netherlands

Liang, X., *Planning and Operation of Automated Taxi Systems*, T2019/13, September 2019, TRAIL Thesis Series, the Netherlands