Relations between the obstacle space of cycling infrastructure and bicycle crashes

An analysis of Amsterdam

Guus van Weelderen

Master thesis
Transport, Infrastructure and Logistics
August 6, 2020

Relations between the obstacle space of cycling infrastructure and bicycle crashes

An analysis of Amsterdam

by

Guus van Weelderen

to obtain the degree of Master of Science
in Transport, Infrastructure and Logistics
at the Delft University of Technology
to be defended on Thursday August 13, 2020 at 11:00 AM.

Student number:	4327381 A.J. van Weelderen	
Project duration:	January 2020 – August 2020	
Thesis committee:	Prof.dr. M.P. Hagenzieker	TU Delft, chair
	Dr. C. Maat	TU Delft, daily supervisor
	Dr. J.A. Annema	TU Delft, daily supervisor
	Dr. G.J. Wijlhuizen	SWOV, external supervisor

This page intentionally left blank.

Preface

During the first seven months of 2020, I have been working hard on this master thesis about relations between the obstacle space of cycling infrastructure and bicycle crashes in Amsterdam, of which the final report now lies in front of you. I have had the pleasure of conducting this study as a graduate intern for the Dutch national institute for road safety research SWOV in The Hague, that made me feel very welcome to be there. While I had a great time at the office, the coronavirus and accompanying measures that were imposed in the Netherlands unfortunately forced me to complete the remainder of the assignment at home after a mere two months. The discovery of certain inaccuracies towards the end even made me rerun a lot of the analyses. Next to the general endeavours of such a project, these *obstacles* were unexpected, but I am glad to have overcome them.

After submitting this report, my time as a student at the Delft University of Technology has almost ended, and I am curious to see what the future will bring. Since 2013, I have first gained a firm basis in the fields of water, construction and transport in the bachelor program in Civil Engineering. Especially the latter made me enthusiastic. The following master program in Transport, Infrastructure and Logistics helped me further pinpoint my passion for anything involving infrastructure and cyclists. However, the various types of this infrastructure, these road users and their vehicles all deserve attention. The *width* of these topics and specialisations is intriguing, and I am fond to have experienced so many of them in order to find my true interests in engineering and be able to contribute to the field.

Before you continue with the rest of the report, I would like to express my gratitude to the people involved. Firstly, I would like to thank my thesis committee: my external supervisor Gert Jan Wijlhuizen for proposing the topic and practically being available every day at the office and online for comments, my supervisors from the university Kees Maat and Jan Anne Annema for their extensive feedback to shape the study and report along the way, and Marjan Hagenzieker for chairing this committee and leading the meetings. Furthermore, I would like to thank Jacques Commandeur for his insights in statistics and modelling, and Dorine Duives for pointing me in the direction of this opportunity after following the Active Modes course. Lastly, I would like to thank my family and friends, for asking me for updates, listening, and shifting my attention every once in a while. The *light* everyone has shed upon the topic and process is very much appreciated.

I hope you enjoy your reading,

Guus van Weelderen
Hoek van Holland, August 2020

This page intentionally left blank.

Summary

Cycling is an important mode of transport in the Netherlands. While the fatality rate of cycling is low (Castro, Kahlmeier, & Gotsch, 2018), the number of traffic deaths, including cyclists, seems to be stagnating after years of reductions (Weijermars, 2019). Like in the Multi-year Road Safety Plan 2016-2021 and Road Safety Strategic Plan 2030, it is important to strive for fewer deaths and injuries (Gemeente Amsterdam, 2016; Kennisnetwerk SPV, 2018). This could be done through safer infrastructure. Previous research already looked into this matter and found that elements such as dedicated cycling facilities, street lighting (Reynolds, Harris, Teschke, Cripton, & Winters, 2009), improvements to the general road safety (Wardlaw, 2014), motor vehicle speed-reducing measures (Schepers, 2008), and lower intersection densities (Wijlhuizen et al., 2016) reduced crash risk for cyclists. However, a relation between the obstacle space of cycling infrastructure and the visibility thereof with bicycle crashes has not been studied well yet. The obstacle space is defined in this study as the space that is available for cyclists to avoid crashes with the obstacles therein. The main question answered in this study thus is the following: *To what extent can the obstacle space of cycling infrastructure and visibility thereof explain bicycle crashes in Amsterdam?* The city of Amsterdam was chosen due to the availability of data from previous studies, its size, its trip modal split and its second place below Copenhagen in a list about bicycle-friendliness of cities (Copenhagenize Design Company, 2020). Sub-questions are related to the general characteristics of the obstacle space of cycling infrastructure, the relation between obstacles on and around cycling infrastructure, cycling infrastructure width, light conditions, and combinations of the aforementioned and bicycle crashes, and measures that can be taken from these relations to improve cycling infrastructure in the Netherlands and reduce crash risk.

From a literature study, it was found that there are multiple obstacle types (Ormel, Klein Wolt, & Den Hertog, 2008), each with their own function and characteristics. Crash risk for those is related to collision, distraction, blocking vision and collision as a result of the latter two. To some extent, these obstacles were included in variables in this study. Suggested widths can vary a lot, depending on the free-space profile around cyclists, use of roads and the intensities on them, with values from 1 metre to 5 metres (CROW, 2018a, 2018b). Less width also allows for less swerving around obstacles and other users, possibly resulting in crashes. Although these guidelines are not compulsory to follow, they are based on experience and evidence, and should not be completely disregarded. The conceptual model in Figure 1 created for this study is an adaptation from literature (Schepers, Hagenzieker, Methorst, Van Wee, & Wegman, 2014) and shows the traffic safety pillars (Othman, Thomson, & Lannér, 2009) that include the obstacle space elements alongside other factors that influence crashes. The influences and interactions of the different parts on each other are illustrated by the arrows. The boxes in blue stand for the focus of this study. Rules and regulations, which were not included in the analysis models, are imperatively considered.

For these relations, data about crashes, exposure (such as volumes) and roadway characteristics are needed. Firstly, for the crashes, ambulance data from the National Institute for Public Health and the Environment (RIVM) and the municipality of Amsterdam from respectively 2009-2012 and 2014-2016 were available. Since not all of these data are logged on exact 4- and 6-digit postal code levels, it was not possible to exactly link certain environments to crashes and the analysis was done on street level. Secondly, the exposure data gives a general idea about how long (street length) or with how many people (intensities) someone is exposed to the variables to be studied, and that may result in a crash. Motor vehicle and cyclist intensities were taken from the traffic model for 2010 by the municipality of Amsterdam and the bicycle intensities are averages of the intensities gathered during the Fietstelweken ("Bicycle Count Weeks") in September in 2015 and 2016. Lastly, the roadway characteristics were extracted from images of distributor roads taken from CycleRAP and captured by Cyclomedia in 2015 and 2016. In total, 337 streets with data were deemed usable.

To answer the research questions, a method is needed that allows for a way to infer bicycle crash numbers from the elements in the conceptual model and the data, and thus streets and the environment. For this, statistical models were applied, specifically generalised linear regression models. This type of model allows for other distributions than the normal distribution to be applied to the dependent variable, such as discrete distributions that are useful for crashes, which are whole numbers (McCullagh & Nelder, 1989). Both the

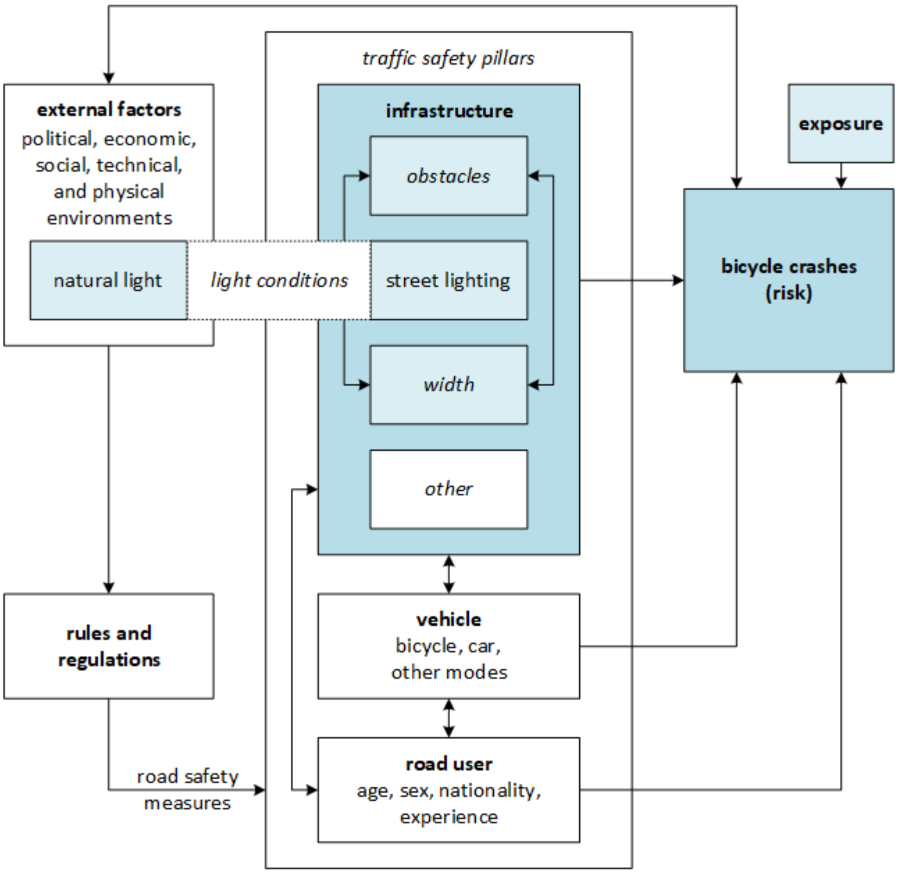


Figure 1: Conceptual model

Poisson and negative binomial distributions were tried. After a check with the Pearson Chi-Square test for the goodness of fit, the negative binomial distribution always performed better.

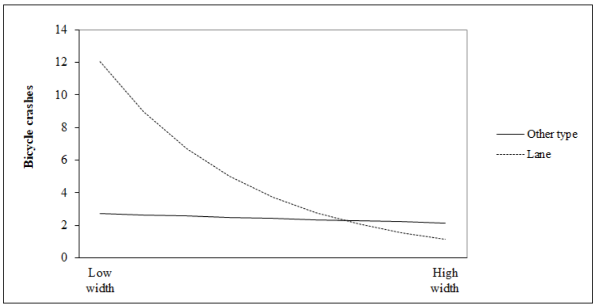
$$\ln(C) = b_0 + \ln(L) + b_1 \ln(I_{mv}) + b_2 \ln(I_{cyc}) + b_3 X_1 + b_4 X_2 + \cdots + b_z X_{z-2} + e \tag{1}$$

Equation 1 shows the general formula used. C are the bicycle crashes on a street, b_0 is the intercept, other b's are regression coefficients, L is the length of a street, I are the traffic intensities (mv for motor vehicles and cyc for cyclists) and X's represent other model variables. The length is an exposure variable, that is generally treated by either putting it in the model as another predictor variable, or by putting it in a (natural) logarithm, with coefficient 1, and adding it as another model variable called the offset. This effectively changes the crash counts to a crash density: the number of crashes per kilometre of road length on a street. Intensities were put in a natural logarithm to simulate a decrease in the rate at which crashes occur at increasing intensities as found in previous research (Wood, Mountain, Connors, Maher, & Ropkins, 2013). After a base model without obstacle space variables, models with separate obstacle space parameters were run and model improvements were checked. This was done on a per-research question basis. The models were extended with interaction variables. As the variables appear in the world together, combined models were developed as well.

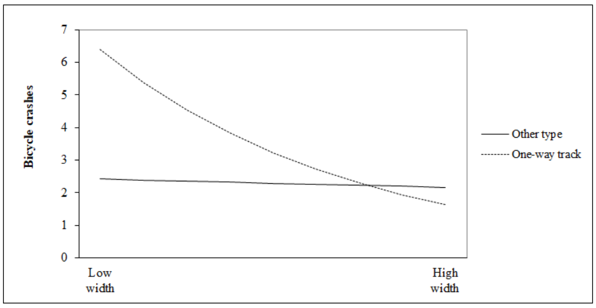
Special operations during the analysis are performed too. Mean centring of all variables is applied, where the means of every independent variable are subtracted from the variables, resulting in means of 0. This is to prevent multicollinearity. Correlations were also taken care of by preventing some combinations from appearing in the models. Conditional variables that can only appear when something is the case (for visibility problems of obstacles) made that main effects of those were not included in interactions, as followed from literature (Dziak & Henry, 2017). Furthermore, since streets typically do not contain only one facility type and we want to look at the impact of facility widths without compromising or negating width differences among streets, it was decided to look at streets with one dominating bicycle facility type where 75% or more was available along the street sections.

Table 1: Analysis results of several crash models

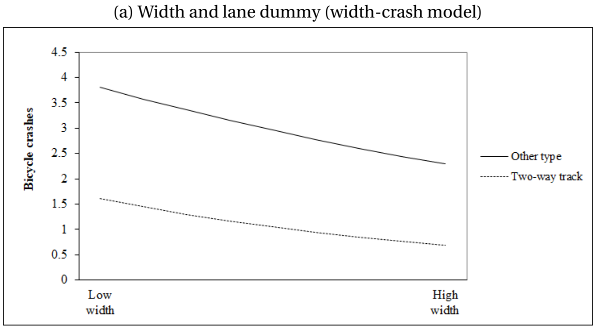
	Model OF		Model WF		Model C2		Model CF	
	PARAMETERS							
	B	Sig	B	Sig	B	Sig	B	Sig
Intercept	-5.55	0.00	-5.50	0.00	-5.57	0.00	-5.53	0.00
Motor vehicle intensity	0.31	0.00	0.31	0.00	0.32	0.00	0.31	0.00
Cyclist intensity	0.75	0.00	0.71	0.00	0.69	0.00	0.71	0.00
Transition bottleneck share	0.35	0.02			-0.16	0.51		
Obstacle-free zone share - 2 m	2.44	0.00			2.46	0.00	1.75	0.00
Dominating width mean			-0.43	0.01	-0.40	0.01	-0.36	0.02
Lane (L) dummy			0.43	0.03	0.46	0.02	0.39	0.04
One-way track (T1) dummy			0.34	0.10	0.18	0.51	0.22	0.29
Two-way track (T2) dummy			-1.04	0.00	-1.10	0.00	-1.07	0.00
Dominating width mean * L dummy			-1.66	0.01	-1.57	0.01	-1.73	0.00
Dominating width mean * T1 dummy			-0.99	0.00	-0.39	0.37	-0.91	0.00
Dominating width mean * T2 dummy			-0.27	0.63	0.39	0.54	0.19	0.75
Dominating width mean * transition					0.66	0.16		
Dominating width mean * OFZ 2 m					-1.99	0.04		
Negative binomial	0.53		0.39		0.36		0.38	
	MEASURES OF FIT							
Pearson Chi-Square / df	1.428		1.345		1.289		1.284	
Log-Likelihood	-799.235		-589.214		-584.730		-584.730	
Akaike Information Criterion (AIC)	1610.470		1200.429		1195.460		1193.460	
Bayesian Information Criterion (BIC)	1633.390		1239.888		1242.094		1236.507	



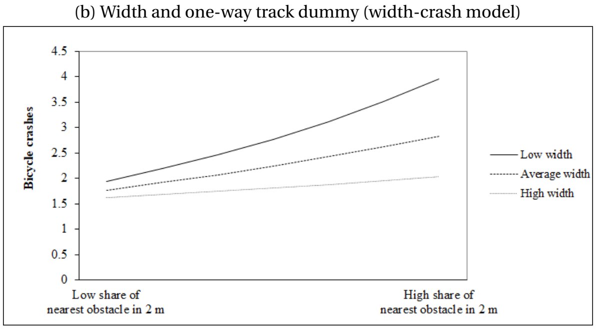
(a) Width and lane dummy (width-crash model)



(b) Width and one-way track dummy (width-crash model)



(c) Width and two-way track dummy (width-crash model)



(d) Width and obstacle-free zone 2 metres (combined model)

Figure 2: Interaction plots

The results of important (final) models are presented in Table 1. Green cells indicate significant parameters, while red cells are insignificant parameters with a p-value larger than 0.05 (parameters are found in column B, p-values in column Sig). Dark grey variables refer to parameters that were expected to contribute to bicycle crashes negatively, meaning positive values will increase the numbers. Light grey cells are the other way around. Since interaction effects are not that easy to interpret, they remain white. Firstly it was confirmed that busier streets lead to more crashes (due to more opportunities for crashes), albeit lowered at higher intensities due to the natural logarithm. Relations between obstacles on and off the road with bicycle crashes were found too. Namely, a higher share of nearest obstacles to the road surface and more bottlenecks in the transition from the road surface to shoulder result in higher bicycle crashes. When combining the scarce on-path obstacles medians and bollards, more visibility problems lead to more crashes (not shown). Relations between width and bicycle crashes appeared as well: wider cycling infrastructure has a positive effect on bicycle crashes (main effect, but also seen in Figure 2c). The width of lanes and one-way bicycle tracks even has an additional effect on bicycle crashes (Figure 2a and 2b). With these results, the hypotheses that more obstacles and narrower infrastructure lead to more crashes were validated. A changing effect between obstacles and crashes was found as well, for an obstacle-free zone at a maximum of 2 metres away from the road surface. The smaller the available space for cyclists is, the higher the crash risk at the same obstacle share (Figure 2d). Apart from visibility problems of on-path obstacles, changing light conditions could not give a reason for crashes in the data set. Overall, it was found that relations between the obstacle space of cycling infrastructure in Amsterdam and registered bicycle crashes in the city exist to a big extent: next to intensities, they were found for the quality of transitions from the road surface to the shoulder, for the share of nearest obstacles in a zone in the shoulder, for the mean width of dominating bicycle facilities and for several interactions between them.

Discussion points and improvements relate to different elements. The light conditions were only briefly touched upon in this study. Furthermore, the effects of the road user and vehicles, that are also part of the conceptual model, were excluded fully in this research due to the focus on infrastructure and the availability of data. Apart from the more general date and time of crashes and in some cases the vehicle type, namely no detailed data about the road user and their vehicle was available. If possible, it would be useful to control for age of cyclists too, since they cycle more during daytime. Considering what has been done in this study, the different types of obstacles in the transition and shoulder can also not be distinguished from the variables, and clustering of obstacles in the obstacle-free zone share is unknown (while this could be more dangerous than simply one obstacle in the shoulder that is closer). As can be derived from the data set, with 42 per cent of the registered vehicles and 45 per cent of those being cyclists, a big share of the crashes had to be disregarded. Registration can be time-consuming at the crash site, underway to the hospital or after treating a crash victim when it is of utmost importance to decrease traffic deaths and injuries. More detailed information still is essential when wanting to make more in-depth studies, so parties that compile this data, such as the ambulance and municipalities, should focus on making the data more complete.

To reduce crash risk, the most straightforward takeaways from this study are of course that cycling infrastructure should come with as little obstacles as possible and be as wide as possible. Problematic medians and bollards on the road regarding visibility lacked profiled road markings and general contrast with the environment. For those obstacles, it is advised to take the guidelines by the technology platform for transport, infrastructure and public space in the Netherlands (called CROW) into account. The quality of the assessed transitions and shoulders depended mainly on the height difference between the road surface and the shoulder. Cracks, bumps and holes are what create these height differences. When these create a high probability of cyclists losing balance, they were marked a bottleneck. Maintenance is therefore of great essence. Lastly, instead of making cyclists use the carriageway, it is better to make any facility like lanes or tracks available. In the case of lanes, this makes sure at least some dedicated space is available, and for tracks, this reduces the exposure to motor vehicles that have to use these facilities. Interestingly, several carriageways in the assessed streets did not have enough width available for all road users to drive next to each other. On streets with speed limits at 50 km/h, speed differences are too high and risk increases. These carriageways also still are frequent on streets. From interactions even followed that an increase in average width on smaller sections of lanes and one-way tracks has a bigger impact on crashes than the same increase on wider parts. It is thus a good idea to take a look at improving smaller sections and moving cyclists from the road to more dedicated infrastructure first.

Concluding, this study contributed more empirical evidence to road safety research, that predominantly focused on matters such as the heavier users of infrastructure: motor vehicles. Especially in other countries, cyclist research and developments lag behind, so it is worthwhile for the Netherlands to take the lead. Relations on street level also help focus on anomalies or peculiarities on detailed sections of streets, that could be hot spots for bicycle crashes. The field of cyclist road safety or the more general active mode road safety is not fully explored yet, and more is certainly still to be done.

Contents

List of figures

List of tables

1 Introduction

1.1. Project context

Cycling is beneficial to both health and the environment, and is relatively cheap (Handy, Van Wee, & Kroesen, 2014). Next to that, it is an important mode of transport in the Netherlands. About a quarter of the trips in the country are made by bicycle, a number which has proven to be quite stable over the years (Buehler & Pucher, 2012). The average exposure in yearly travelled kilometres per person for cycling even is among the highest in the world, and per hundred million kilometres cycled, 0.8 fatalities are registered. When looking at the five-year averages of cyclists deaths in traffic for several European countries (Figure 1.1), a higher exposure for cycling per person per year in a country suggests a lower risk of getting killed in a crash. Since the Netherlands places itself at the right-hand side of the graph, the country thus already performs relatively well in this aspect (Castro et al., 2018). Still, the risk per kilometre travelled to die in traffic in 2009 was about a factor 5 higher for cyclists than for someone in a car, and to get severely injured about 40 (SWOV, 2012). As will be shown later, the trend in traffic deaths is far from promising, and cyclist involvement in crashes is rather high. It thus is worthwhile to look for different ways to make cycling safer and reduce the involved risks.

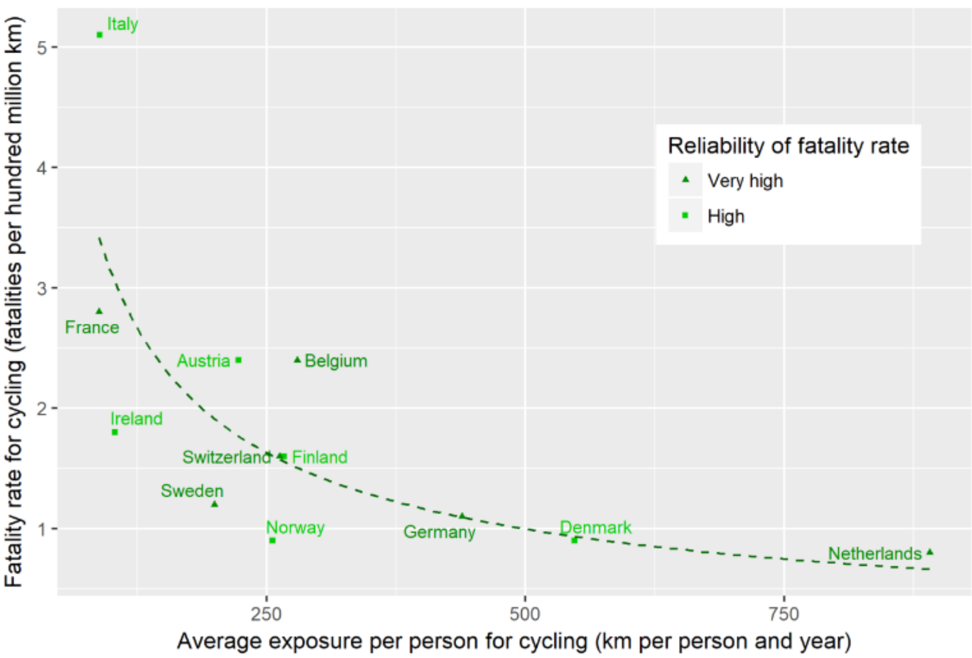


Figure 1.1: Fatality rate versus exposure for cycling (Castro et al., 2018)

1.2. Research problem

One way to make cycling safer is by focusing on the infrastructure. Various implementations of the actual road itself, with or without facilities just for cyclists and accompanying markings, traffic signs and lights are examples of this infrastructure. But for this to work, general relations between infrastructure and bicycle crashes need to be known: which elements improve safety, and which decrease it? This study focuses on the obstacle space for cyclists, which was defined in this study as the space that is available for cyclists to avoid crashes with the obstacles therein, and the visibility of this space. Infrastructure and its relation to crashes have been researched numerous times in the past, but the aforementioned does not seem to be part of that. The following paragraphs try to back this statement up. Crashes that occur on this infrastructure can be split into single-bicycle crashes and other crashes (such as bicycle-motor vehicle crashes). If one of the following studies addressed any of these in particular, this will be made clear.

In 2009, a review paper that examined 23 different papers found that facilities that are specifically built for cyclists reduce crashes and injuries for cyclists (Reynolds et al., 2009). To some extent, this was acknowledged by a more recent paper that also looked at the impacts of cycling infrastructure on safety: of the 22 bicycle treatments that were explored, only bicycle streets (priority for cyclists over motorised vehicles) and bicycle tracks (separated from roads) were found to reduce crash risk. Moreover, cycling at higher volumes appeared to positively influence safety. This is known as the "safety in numbers" phenomenon (DiGioia, Watkins, Xu, Rodgers, & Guensler, 2017). Next to sufficient street lighting, paved surfaces and lower-sloped roads (Reynolds et al., 2009), improvements to the general road safety are beneficial as well (Wardlaw, 2014). Bicycle-motor vehicle crashes in urban areas in the United States furthermore were found to happen less at higher carriageway lane widths, at higher speed limits and with grass in the median, while sidewalks and sidewalk barriers increased them, among several other things (Raihan, Alluri, Wu, & Gan, 2019). This last study will also be mentioned briefly in the methodology.

Specifically looking at the Netherlands, Schepers (2008) studied cyclist crashes and infrastructure a lot. The risk of single-bicycle crashes due to infrastructure was found not to differ significantly between road types (carriageways, bicycle tracks, bicycle lanes, etc.). Furthermore, about half of the single-bicycle crashes were caused by infrastructure-related factors. The influence of the width of bicycle lanes was mentioned as a future research possibility. Later studies (such as Schepers and den Brinker (2011) and Schepers and Klein Wolt (2012)) examined the influence of the road design and network characteristics on road safety for cyclists. Its main focus was on bicycle-motor vehicle crashes (at unsignalised priority intersections) and single-bicycle crashes. Again, the road design appeared to influence cyclist safety. Infrastructure-related crashes were caused by collisions with obstacles, riding off the road, skidding bicycles due to slippery road surfaces and the inability to stabilise bicycles or stay on the bicycle due to uneven road surfaces. Other causes were loss of control (low speed, forces on the front wheel and poor or risky riding behaviour), bicycle defects and falling due to an external force. In situations where cyclists had right of way, on intersections with two-way bicycle tracks that were marked and coloured red, crashes happened more. On the other hand, raised bicycle crossings and other measures that aim to reduce the speed of motor vehicles reduced the risk. One of the main recommendations of the report was to continue researching the relation between characteristics of road infrastructure and road safety.

A study by Wijlhuizen et al. (2016) created a compound obstacle score, but no relation could be found with bicycle crashes due to few obstacles along the captured roads. (This relation was present in Wijlhuizen et al. (2017), but it did not only address cyclists and gradually changed the method used earlier.) No relation between the quality of cycling infrastructure, including width, was found either. They thus state it should be investigated in other locations. Several other relations could be found with the number of bicycle crashes per kilometre of road length, however. For example, higher cycling and motor vehicle intensities, and higher densities of intersections both relate to a higher number of bicycle crashes. The severity of injuries did not seem to make a difference here, so it was recommended to only use one model with just the general crash numbers.

As it is assumed that light has an impact on crashes, the effect of light conditions has also been a research topic. Wanvik (2009) looked at the effect of road lighting on crashes that happened in twilight and darkness in the Netherlands in a period of almost 30 years. Ratios for lit and unlit roads in daylight and darkness for several hours of the day were used, thus excluding any differences in intensities. Overall, road lighting was found to decrease crashes with injuries by 50% during darkness. During twilight, the effect was only 2/3 of that. In rural areas, lighting seemed more effective. Road lighting also had more effect on pedestrians, cyclists and moped users compared to other motor vehicles. While the paper looked at different situations, such as weather conditions, road users and collision types, it did not take into account infrastructure characteristics.

Schepers and den Brinker (2011) furthermore specifically studied visibility and visual design on single-bicycle crash sites, where light conditions also were a variable. The research used a questionnaire study to find the number of occurrences (incidence) of certain factors involving the accident itself, and a special method to computationally investigate the visual characteristics of several crash sites of single-bicycle crashes. In the first of two studies, crashes seemed more frequent under conditions with less light, but this was not significant – it is suggested that this is due to older cyclists not cycling much in darkness (see Figure 1.2). Collisions with bollards or narrowings or riding off the road were found to be the result of bad visual characteristics, recommended to be prevented by edge markings and better visible bollards. They also propose looking into whether motorists would use cycle tracks if bollards would be removed, and limiting consequences of bicycle crashes by designing bicycle facilities wide enough, with a small height difference with the sides. Suggestions for improvement of visibility were previously given in Fabriek, De Waard, and Schepers (2012) as well.

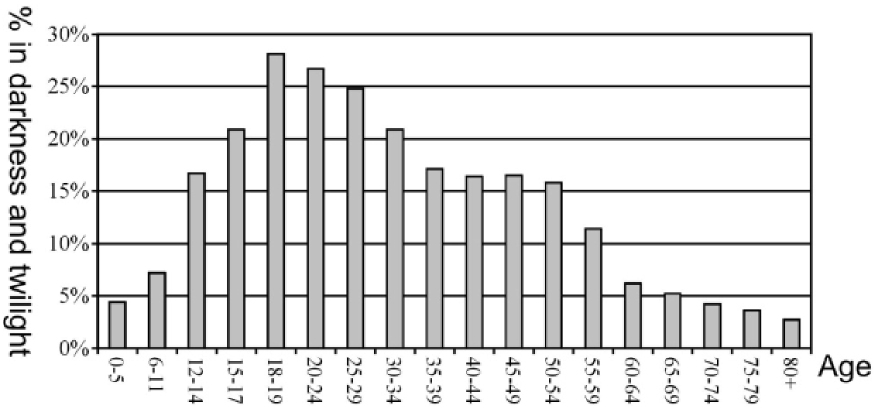


Figure 1.2: Share of distance travelled by bicycle in darkness and twilight per age group Schepers and den Brinker (2011)

The quality of infrastructure elements, or design in general, is mentioned in Schepers et al. (2014) as an important point; not merely the presence of elements is of interest, but also how they are implemented. Differences between countries furthermore make that the aforementioned risk-reducing or increasing factors may not hold everywhere.

From the above paragraphs, it can be concluded that research into behavioural and infrastructural risk factors with regards to cyclists has been done in various ways in the past. With cyclists and single-bicycle crashes taking up the highest shares in traffic deaths and severe injuries, and infrastructure being the most common cause of single-bicycle crashes, its importance is also clear. However, a research gap indeed still exists. None of the studies seem to have looked specifically into the crash risk of obstacles for cyclists and how this risk changes in combination with different width and light conditions, basically meaning the ease at which these obstacles can be avoided in the form of space and visibility constraints. When cyclists are riding their bicycle and face an obstacle, the width available namely limits their manoeuvring space. If the road section they ride on is several metres wide, there is a higher chance they can swerve around it or otherwise avoid a collision without loss of balance than when it is only half a metre wide. The same holds for light conditions: during the day, it is easier to see obstacles, but as light decreases, so does the ability of cyclists and other road users to notice obstacles. In conjunction, say obstacles on a route with narrower road sections during the night, it can be much harder to avoid a crash. This gap will thus be addressed in this research, although light conditions will only be briefly touched upon due to scoping later on in the process.

1.3. Research objective

To test if any of the previously stated relations between the obstacle space and bicycle crashes are present, a real-world application is empirically tested. The Dutch national institute for road safety research SWOV (see Appendix B for more information about the institute), has conducted studies into infrastructure and road crashes involving both all modalities and cyclists in the recent past, partly commissioned by the municipality of Amsterdam (Wijlhuizen et al., 2016, 2017). As such, data are available and the studies can be supplemented with new results. With the city being the capital and biggest city in the country that receives a lot of tourists, a lot of cyclists make use of the infrastructure. Of the four biggest cities in the Netherlands (Amsterdam, Rotterdam, The Hague and Utrecht), Amsterdam comes second when looking at the trip modal split, with 33% being made by bicycles (Schaap, Harms, Kansen, & Wüst, 2015). First is Utrecht, although it ranked below Amsterdam in the latest Copenhagenize Index from 2019. This index ranks cities in the world according to their bicycle-friendliness. Amsterdam places second in this list as well, below Copenhagen (Copenhagenize Design Company, 2020). The city of Amsterdam is thus deemed a good case study again.

The objective of this study is to address the gap in literature of the relation between obstacles, width, light conditions and their interactions with bicycle crashes in Amsterdam through structured methods and with (pre-)gathered data, as to help researchers, municipalities and other organisations understand more about cyclists and cycling and help the Netherlands become safer and more bicycle-friendly. For example, any relations found can be put into practice by adjusting unsafe infrastructure on bicycle routes and used as further evidence for CycleRAP.

1.4. Research questions

To structure the research, the following research questions have been defined. As mentioned, the main question aims to answer whether separate infrastructure elements related to obstacles and width can predict bicycle crashes, also taking visibility thereof in the form of light conditions into account. Furthermore, it is relevant to know whether a variable influences the outcome differently when another changes. The sub-questions expand more on these points. The modelling questions are preceded by a literature question to better interpret the elements and propose sensible measures with the last question that aims to make sensible recommendations regarding available space for swerving and visibility.

Main question:

To what extent can the obstacle space of cycling infrastructure and visibility thereof explain bicycle crashes in Amsterdam?

Sub-questions:

1. What are the general characteristics of the obstacle space of cycling infrastructure?
   (a) What are the types of obstacles and range of widths?
   (b) What are the determining factors for implementation?
   (c) What are the determining factors for crash risk?
2. What is the relation between obstacles on and around cycling infrastructure and bicycle crashes?
   ⟶ *Hypothesis: more obstacles lead to more crashes*
3. What is the relation between cycling infrastructure width and bicycle crashes?
   ⟶ *Hypothesis: narrower infrastructure leads to more crashes*
4. What is the relation between light conditions and bicycle crashes?
   ⟶ *Hypothesis: darker conditions lead to more crashes*
5. What is the relation between combinations of the above factors and bicycle crashes?
6. What measures can be taken from these relations to improve cycling infrastructure in the Netherlands and reduce crash risk?

1.5. Reading guide

The previous questions will be answered throughout the report. Next to more background information about crashes in traffic, Chapter 2 will discuss the first question about obstacles and width, and end with a conceptual model. In Chapter 3, information about data gathering and processing and the general methodology for the analysis will be explained. Chapter 4 consecutively elaborates on the data and which variables are created, after which the next set of questions are answered by taking a look at the modelling results in Chapter 5. In the final chapter, the study is concluded, some shortcomings and further steps are discussed and recommendations are made.

2
Literature research

While the introduction already contained literature research to get acquainted with the topic of this study, such as the gap, more research was needed. The first question was formulated specifically for that. After a closer look at trends in crashes and the share cyclists take in this, the sections thereafter will cover characteristics of obstacles on and around cycling infrastructure and the accompanying widths of that infrastructure, specifically the types, determining factors for implementation and crash risk. The chapter ends with the conceptual model, that contains these parts of the infrastructure, that were also investigated in previous studies.

For the literature review and analyses later on, peer-reviewed articles, papers and research reports were acquired from online databases like Scopus, ScienceDirect, ResearchGate, Google Scholar and simple search engines. Being a student at the Delft University of Technology, getting access to papers proved to be easier. Reports and fact sheets from SWOV were used as well, that are available internally and online. Keywords and combinations of keywords used for the search of relevant material included (in alphabetical order) *crashes, cycling, obstacles, infrastructure, light conditions, relation, traffic, width* and others, including synonyms. Given that research into cycling is rather new, newer research was favoured, preferably from 2000 onward. Experts and practitioners in the field from SWOV and the university gave valuable contributions too if not all information was available online or could not be found, and provided certain papers or helped structure methods if needed.

2.1. Traffic deaths and injuries

Of the 123,000 traffic victims that were treated on the emergency department in 2018, just over sixty per cent were cyclists. The number of cyclists that was treated for severe injuries even increased to 46,800, which is nearly a third higher than in 2009 (VeiligheidNL, 2019). In the period from 1996 to 2014, the number of cyclist deaths without interference from motor vehicles increased by 7% per year, while those with motor vehicles decreased by 3.8% per year. Despite that, the total number of cyclist deaths seemed to stay stable during that period, while the annual road deaths through other modes more than halved (Schepers, Stipdonk, Methorst, & Olivier, 2017). The total trend appears to have stopped in recent years, however, going up again from 2013 (orange points in Figure 2.1), with 65 more traffic deaths in 2018 than in 2017 (Weijermars, 2019).

Of the total traffic deaths and severely injured by traffic, cyclists make out a big part. Figure 2.2 shows the road casualties in the Netherlands in 2018. In orange (dark and light), cyclist shares and numbers are given. 228 of the deaths were cyclists, and 52% of these cyclist deaths are 60 years or older. The numbers of deaths and injured are generally thought to be subject to developments in mobility, and ageing of the population in the last ten years is given as a partial explanation for stopping the downward trend. Another notable change that could have had an influence is the change in the vehicle fleet: the number of mopeds and electric bicycles has increased a lot. Registered mopeds almost doubled from 2009 to 2019 (to 750,000) and 40% of the sold bicycles in 2018 was electric, with just over 400,000 electric bicycles (Weijermars, 2019). Based on a continuation of the current transport policies, a forecast for the Netherlands assumes 38% of the traffic deaths and 62% of severely injured through traffic to be a cyclist in 2030 (Weijermars et al., 2018). These are comparable to the current shares.

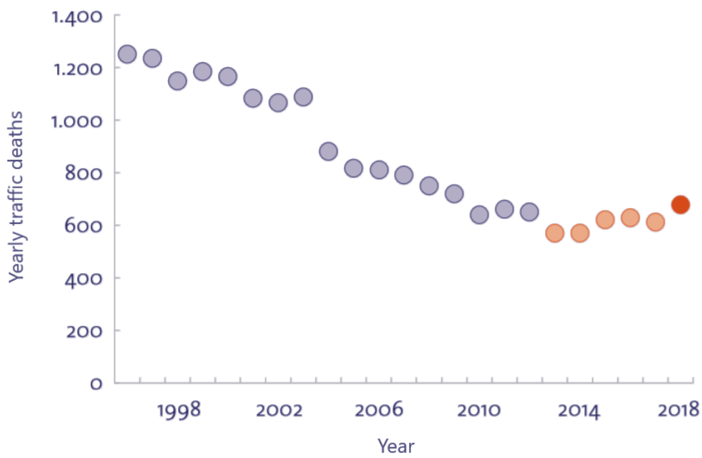


Figure 2.1: Yearly traffic deaths (translated; taken from Weijermars (2019) with data from CBS)

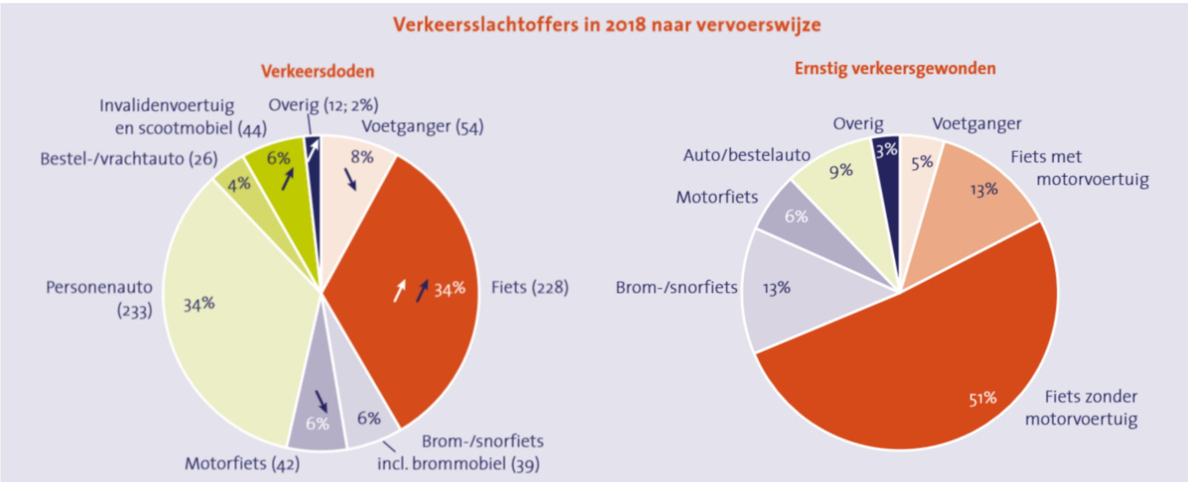


Figure 2.2: Share of traffic deaths (left) and registered severely injured by traffic (right). Blue arrows indicate significant developments over the last ten years, and white arrows 2018 compared to 2015-2017 (Weijermars, 2019). Cyclists are coloured in dark and light orange.

Since it is not feasible for the upward or even stagnating trend of traffic deaths to continue, and cyclists to be such a big part of that, things need to change. A reduction in the number of traffic deaths is already part of several plans. Firstly, there's the Strategisch Plan Verkeersveiligheid 2030 ("Road Safety Strategic Plan 2030") by the Dutch government, provinces, municipalities, vervoerregio's ("traffic regions") and societal organisations. The vision has the ambition to achieve zero traffic deaths in 2050, just like that of the European Commission. In the report, several policy themes are described that pose a risk on traffic safety, of which safe infrastructure is one, with possible solution directions (Kennisnetwerk SPV, 2018). Moreover, a first fact sheet gives five of the most effective and supported measures, such as separating cyclists from motorised traffic on 50 and 80 km/h roads and making cyclist infrastructure safer, for example through side markings and a smoother road surface (Kennisnetwerk SPV, 2019). The municipality of Amsterdam has a plan of its own as well, the Meerjarenplan Verkeersveiligheid 2016-2021 ("Multi-year Road Safety Plan 2016-2021"). In it, they stress the increasing crowdedness in the city and the effects on the infrastructure. Measures such as tackling so-called black spots, locations where multiple crashes happen, and listing the state of the current infrastructure in order to preventively improve it (i.e. through the Network Safety Index) are part of this plan. Several traffic participants are specifically targeted, of which one group is cyclists. Examples of measures taken to reduce single-bicycle crashes (crashes where cyclists are the only participant in a crash) are side markings and flexible bollards (Gemeente Amsterdam, 2016).

A proactive method, that tries to act before crashes happen, is a recent project by SWOV called CycleRAP, formerly known as Safe Cycling Network. It is a method to score cycling infrastructure based on the probability of a crash (1 star equals the highest risk, 5 stars the lowest), that accounts for individual infrastructure elements in order to see which roads can be improved (Wijlhuizen et al., 2016). It was commissioned by the Dutch travellers' association in 2013. The method is part of the larger international Road Assessment Programme iRAP (ANWB, n.d.), and of the earlier mentioned Network Safety Index. The initiative arose due to the need from "road managers" ("wegbeheerders") to map cycling safety, and the association's involvement in the Decade of Action for Road Safety 2011-2020 by the UN General Assembly, that strives to save lots of lives worldwide through better management of road safety measures, such as the infrastructure itself, but also vehicles and behaviour (World Health Organization, 2019). As mentioned, this study hopes to contribute to these projects and goals.

2.2. Element characteristics

To interpret the results of the coming analyses, it is useful to grasp the underlying elements related to obstacle avoidance and how these can lead to crashes. The sub-question "What are the general characteristics of the obstacle space of cycling infrastructure?" was defined for this purpose. In the following sections, a look will be taken at the types of obstacles and ranges of widths, and determining factors for implementation and crash risk.

In several cases, the guidelines by CROW, the technology platform for transport, infrastructure and public space in the Netherlands, are cited. The guidelines for cycling were first drafted in the nineties, based on experience from workgroups based on consensus, logical reasoning and simply adding measurements, and evidence from previous studies. Bicycle policies are often local, which means the local situation needs to be taken into account. Therefore, the guidelines are supplemented with arguments, experiences, ideas and tips (CROW, 2018b). Next to that, the guidelines do not have a legal status, so it is allowed to deviate from them. In a study among several municipalities, 56% of the respondents mention they barely use the Ontwerpwijzer Fietsverkeer (Design Manual for Bicycle Traffic), if at all. After assessing the implementation in two municipalities, mostly the recommendations for the width of bicycle tracks, obstacle-free zone and profiled road markings at bollards are not followed. Lack of space, costs and awareness of the guidelines are reasons why the guidelines are not always applied (Bax, van Petegem, & Giesen, 2014).

2.2.1. Obstacles

An obstacle in the most general sense is "something that blocks you so that movement, going forward, or action is prevented or made more difficult" (Cambridge University Press, 2020). In the event of a (near-)collision with an obstacle during cycling, this can lead to a crash for cyclists and/or other road users.

A first distinction that can be made is found in Schepers and Klein Wolt (2012). Based on earlier research, they proposed several crash categorisations. The crashes related to the infrastructure were split over those caused by the cyclist straying from their regular path unintentionally or linked to the road surface quality. The first was split further in colliding with either obstacles on the road that were put there by road authorities, and parked vehicles, or riding off the road and colliding with obstacles there. The road surface quality was split to skidding due to a slippery road surface and loss of control due to an uneven road surface or a loose object. The road surface quality by itself is not thought to be an obstacle, but defects there can lead to collisions with obstacles on or off the road. Derived from this, the distinction is thus whether it is located on or off the road (the "site"). Further distinctions that can be made are whether the obstacle is fixed or dynamic, i.e. is it always there or only occasionally (the "state"), and whether it is a point or line obstacle, i.e. is it concentrated on a single spot or a longer stretch (the "shape")?

Typical obstacles considered in a survey by Ormel et al. (2008) and used by Schepers and Klein Wolt (2012) are lamp posts, traffic signs, bollards, fences, walls, kerbs and trees, along with the option "Other" and animals which are not infrastructure-related and thus not considered here. Under "Other", a type of obstacles to consider are parked cars on parking places. Hereafter, these obstacles are all briefly elaborated. Next to the general collision risk, other possible crash risks are mentioned.

- Lamp posts provide lighting to the public, that has multiple functions, of which traffic safety and social security are the most important. Other functions, such as providing a pleasant living environment, are more related to how people experience public spaces. Related to traffic safety, light makes the road profile and sides of the road, information provided by for example traffic signs and other road users more visible, thus possibly reducing crash risk. On the other hand, too much light can work blinding, reduce contrast, distract and more, and should be avoided (CROW, 2017). Since they emit light for an area, multiple in the same spot are not needed, and there is space between them. They are fixed point obstacles, placed next to the road or on medians.

- Traffic signs come in many forms and all have the purpose of providing information to road users. They are often placed on poles, sometimes with multiple signs at once. This can be distracting. Earlier research states that 20% of the signs in the Netherlands can even be removed (Kerssies, n.d.). They are fixed point obstacles, placed next to the road or on medians.
- Bollards are posts that are usually put on or next to the road to prevent motorised traffic to enter a road stretch. Since they are obstacles, can narrow the path for road users and are cumbersome on routes for deicing of the roads, CROW (2018b) suggests only placing them when other measures are not sufficient and placement outweigh these disadvantages. This should be done in a contrasting colour palette, with a profiled road marking and sufficient lighting. Bollards are fixed point obstacles, but multiple can be put in a row. Still, some space between them should be present.
- Fences and walls are fixed line obstacles that are found next to the road. Fences can be placed on medians to create a staggered crossing, where road users face oncoming traffic when they cross a road. They can guide pedestrians and road users to intended crossing points, and prevent parking on the pavement (iRAP, 2010). They are a barrier that can guard, but may also make it harder to avoid a crash since it blocks one side of the road. Walls furthermore are built to support structures. Depending on the height and ability to see through the fence or wall, vision can be blocked.
- Kerbs are "raised edges along the side of a street, often forming part of a path for people to walk on" (Cambridge University Press, 2020). In the light of traffic safety, they define the limit of the carriageway and walkway, and as such form a barrier between vehicles and pedestrians and control access. They furthermore allow for controlling the drainage and reducing erosion of the road. There are vertical and sloped kerbs, with the latter designed in a way they are more easily passable (AASHTO, 2018). This is useful in case a road user needs to swerve. Kerbs are fixed line obstacles that are mostly found next to the carriageway, but are also used for medians.
- Trees and other vegetation can have a positive impact on life in a city and humans in general in many ways (Burden (2006) for example mentions 22 benefits), but they also have proved to positively influence traffic safety (Kocur-Bera & Dudzinska, 2015). They can, however, block the vision of areas behind it. Trees are fixed point obstacles next to the road. Other vegetation can also be present in a line.
- Parking places and accompanying parking and parked cars lead to extra movements and obstacles on and next to the road. Cars are not always present in parking places, however.

An overview of all functions and the classification as written in the sections before is shown in Table 2.1.

Table 2.1: Overview obstacles with function and classification

Obstacle	**Function**	**Classification**		
		Site	*State*	*Shape*
Lamp post	Provide light	Off	Fixed	Point
Traffic sign	Provide information	Off	Fixed	Point
Bollard	Control access for motor vehicles	On/off	Fixed	Point/line
Fence	Define barrier, control access and guard	Off	Fixed	Line
Wall	Define barrier, support structures and guard	Off	Fixed	Line
Kerb	Define barrier, control access and more	On/off	Fixed	Line
Tree/vegetation	Define barrier, reduce speeds and more	Off	Fixed	Point/line
Parking place	Store vehicle	On/off	Dynamic	Line

2.2.2. Width

The width of a road for cyclists depends on several factors. Firstly, it is important to take the dimensions of the users and any vehicles they use into account. As cyclists and bicycles come in many forms, it is not useful to design infrastructure for the average bicycle, and the legal requirements serve as a better basis. These state that a bicycle can be 0.75 metres wide at maximum, and mopeds 1.10 metres. Next to this, some extra space is added. Since a bicycle just has two wheels, the rider namely needs to keep their balance and as a result makes small corrections while cycling. This results in a sway ("vetergang" in Dutch), that makes that cyclists do not ride in a perfectly straight line. These come together in the free-space profile ("profiel van vrije ruimte"), that is shown in Figure 2.3 and is the space around a cyclist a designer needs to bear in mind (CROW, 2018b), similar to the loading gauge of trains. Another factor to take into account for the width of paving is the fear of obstacles on the sides, including other cyclists, with a safety distance of at least 0.25 metres. With every cyclist in a cross-section that gets added, the width ideally thus needs to be increased by at minimum 1 metre. Furthermore, the intended use (either one or two directions, and possibilities for overtaking) and the expected traffic intensities are of interest. For a bicycle-moped track, 0.5 metres is usually added (CROW, 2018a). Depending on a combination of the aforementioned, the minimum width values per direction for cyclists typically range from 2 to 5 metres, with a special mention for 1.5 metres for solitary bicycle tracks at low intensities and with easily passable kerbs or shoulders and at least 1 metre on the carriageway.

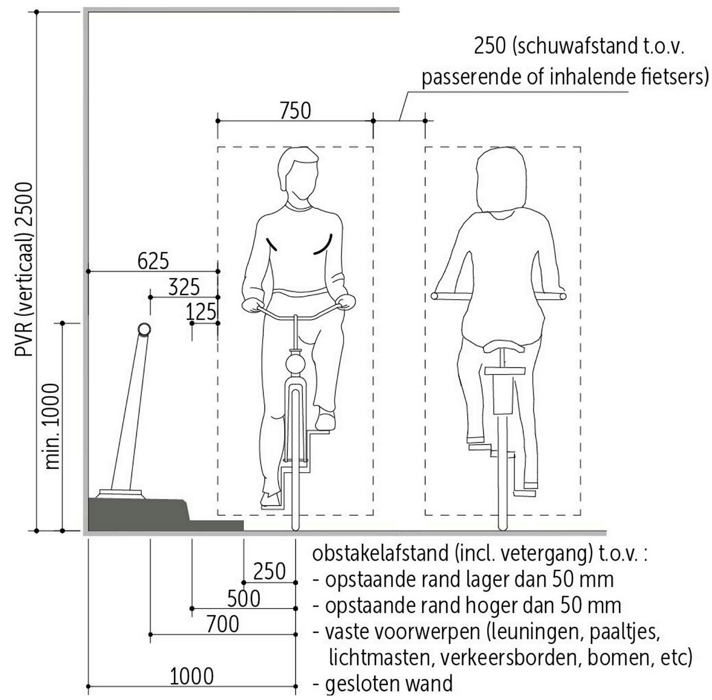


Figure 2.3: Free-space profile with distances for cyclist on bicycle (CROW, 2018b)

These widths express themselves in different ways, depending on where they are available. Next to the main carriageway that does not have any dedicated cycling infrastructure, cyclists in the Netherlands have access to a bicycle track, bicycle-moped track, non-designated bicycle lane, bicycle lane and bicycle street (a high-quality bicycle connection, that allows for shared use by motor vehicles). Section Fietsvoorziening of Appendix C shows some visual examples of these facilities. Considerations that need to be taken into account for the carriageway are that smaller profiles can contribute to lower speeds, that lower intensities make that these are relatively generous and that a lot of parking movements might increase unsafety. A non-designated bicycle lane usually is a downgraded bicycle lane without a bicycle symbol on it, and does not have a legal status. The facilities all have the function to offer a connection for either cyclists, cyclists and mopeds or all sorts of vehicles. Bicycle lanes additionally serve to indicate and secure the position of cyclists on a road section, and bicycle tracks exist to separate motorised and cyclist traffic for the safety and comfort of cyclists (CROW, 2018b).

Several of the previous factors can also be derived from Table 2.2, that gives suggestions for the provision of the facilities in the built-up area. A division is made by road category, speed limit, intensity and bicycle network category. An access road gives access to origins and destinations. Both on the road sections and intersections, exchanges take place. On the other hand, on a crossing of distributor roads exchanges only take place on intersections, where access roads are connected to other access roads and through-roads (motorways and trunk roads) (CROW, 2018c; SWOV, 2017). For example, bicycle tracks are recommended in areas with access roads with a higher intensity and distributor roads, and a higher speed limit (and thus higher speed differences).

Table 2.2: Suggested bicycle facilities inside built-up area (CROW, 2018b)

<table>
<tr><th colspan="4"></th><th colspan="3">Bicycle network category</th></tr>
<tr><th>Road category</th><th colspan="2">Maximum speed of motor vehicles (km/h)</th><th>Intensity of motor vehicles (m.v./day)</th><th>Basic infrastructure (<750 cycl./day)</th><th>Main bicycle network (500-2,500 cycl./day)</th><th>Quick bicycle route (2,000 cycl./day)</th></tr>
<tr><td rowspan="3">Access road</td><td colspan="2" rowspan="3">Walking pace or 30</td><td><2,500</td><td rowspan="2">Mixed traffic</td><td>Mixed traffic or bicycle street</td><td>Bicycle street (with priority)</td></tr>
<tr><td>2,000 - 5,000</td><td>Mixed traffic or bicycle lane</td><td rowspan="2">Bicycle track/lane (with priority)</td></tr>
<tr><td>>4,000</td><td colspan="2">Bicycle lane or bicycle track</td></tr>
<tr><td rowspan="3">Distributor road</td><td rowspan="2">50</td><td>2x1 lanes</td><td rowspan="3">Not relevant</td><td></td><td colspan="2" rowspan="2">Bicycle track</td></tr>
<tr><td>2x2 lanes</td><td></td></tr>
<tr><td>70</td><td></td><td colspan="3">Bicycle-moped track</td></tr>
</table>

2.3. Conceptual model

To prepare for the methodology, the previous points should be part of a greater scheme that shows the hypothesised relations: a conceptual model. The basis for this can be taken from literature about traffic safety. A road safety framework was constructed by Schepers et al. (2014), that consists of travel behaviour, exposure to risk and risk itself, leading to crashes and injuries, and links several elements with theories. For this, the framework as shown in Figure 2.4 combines the passenger transport model by Annema and van Wee (2009) with the traffic safety pillars by Othman et al. (2009), namely human (road user), vehicle and infrastructure. As mentioned previously, along with external factors, these are also the categories of single-bicycle crash causes found by Schepers and Klein Wolt (2012).

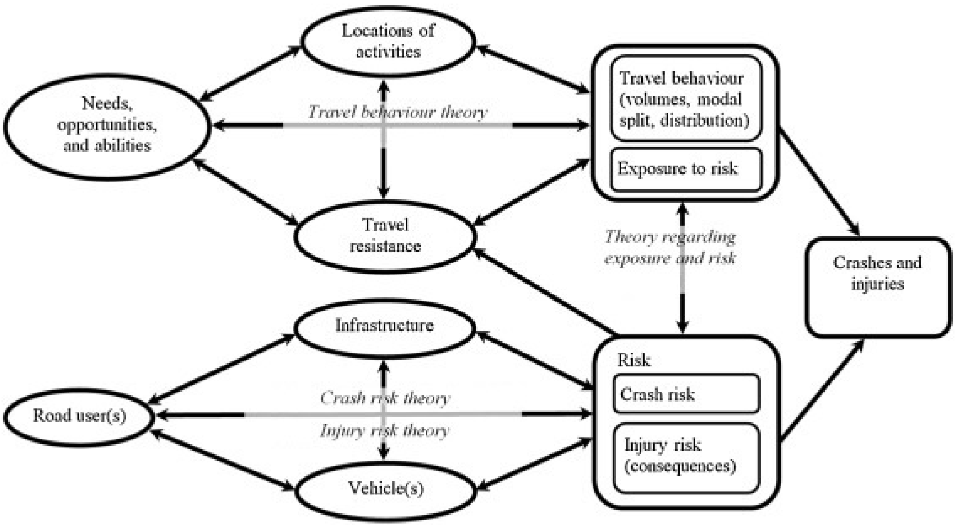


Figure 2.4: Conceptual framework for road safety, including exposure and risk (Schepers et al., 2014)

Since this framework is quite extensive in itself and not all parts are needed, the conceptual model as depicted in Figure 2.5 was developed, based mainly on the bottom part of the previous one. As can be seen, the three pillars and crashes still take up an important place. The influences and interactions of the different parts on each other are illustrated by the arrows. The boxes in blue stand for the focus of this study. Rules and regulations, which were not included in the analysis models, are imperatively considered at the end.

The different road users, here solely those involved in bicycle crashes (thus cyclists themselves and possibly motorists) each have their own capabilities and make decisions that influence the risk, either on purpose or not. Some people choose to take more risk, for example by driving faster than allowed. Characteristics such as age, sex, nationality and experience are assumed to be of influence on these risks. The whole psychological and biological aspect of humans can be of great interest for behaviour and research into crashes, but since the scope of this project limits itself around infrastructure, only their presence on the roads is used.

Vehicles influence the risk as well. These can differ per type and number of vehicles that drive on a road, but

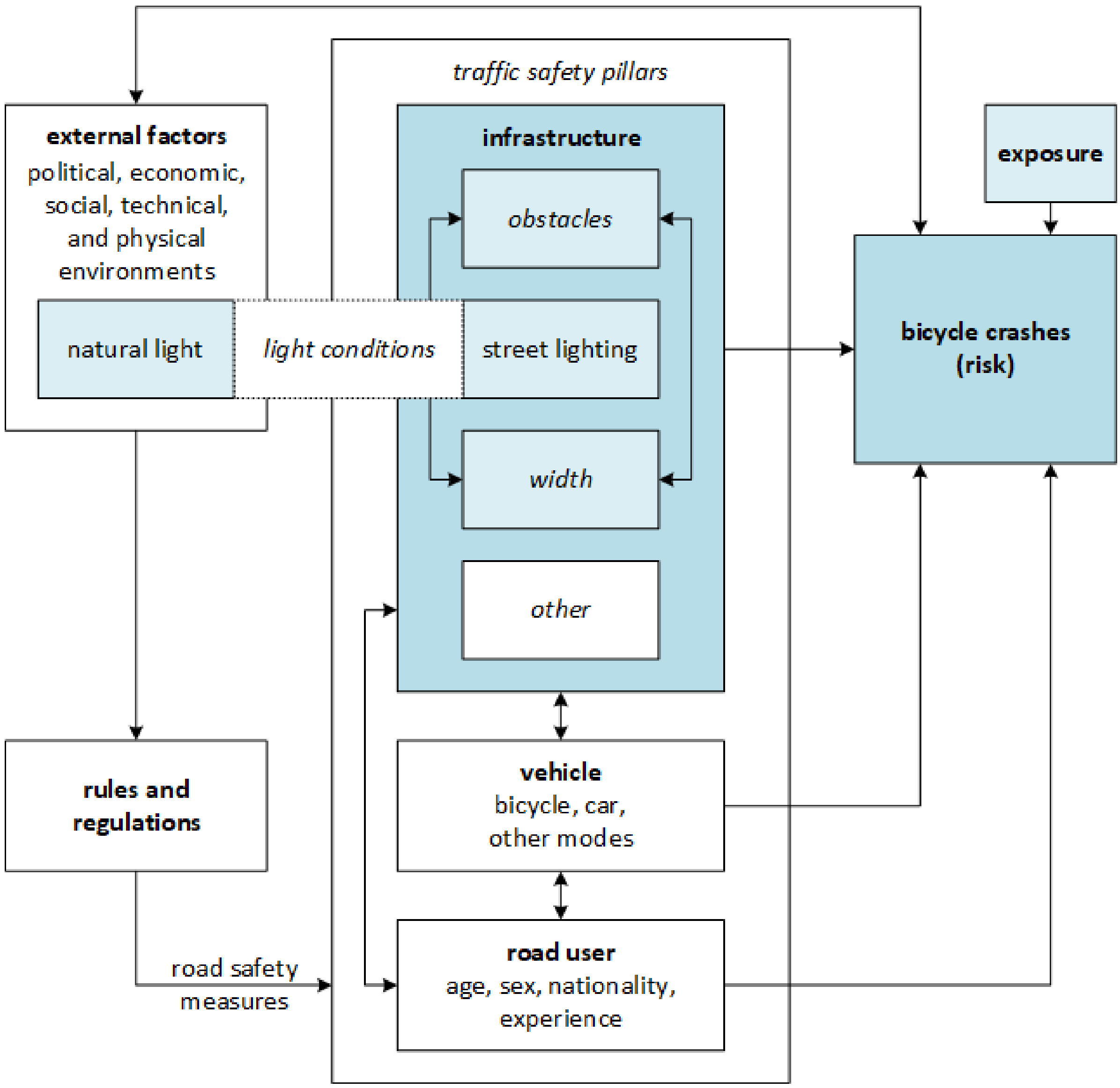


Figure 2.5: Conceptual model

also the state they are in. A bicycle with missing lights or a car that did not go through periodic maintenance probably has a higher crash risks than vehicles that do have or did the aforementioned. While it would be valuable to also test for data surrounding vehicle, these are not present and will thus not be taken into account, apart from the fact that a bicycle needs to have been involved for a crash to be called a bicycle crash.

As defined before, the infrastructure comes in different forms, such as roads and intersections, with entries and exits, various signs, different widths, and more. For this study, the obstacles, width and street lighting are of main importance. Any other factors that are not related to those factors later on are in the category "Other". Internally, these elements might affect each other, such as the width and obstacles. The characteristics of the infrastructure that were categorised and put in the data set can be found in Appendix C.

Both direct and indirect effects of the infrastructure characteristics are expected. If everything else remains the same, the hypotheses are the following. More obstacles on and next to the road increase the risk of colliding with one. Furthermore, the wider the bicycle road infrastructure, the more space is available for cyclists to avoid crashes, for example due to collisions with obstacles. With more natural light and street lighting available, people can feel safer and can be less afraid to ride faster. As such, more crashes may happen. Indirect relations with bicycle crashes are assumed to be between these light conditions, width and obstacles. If a street has a certain number of obstacles but has better light conditions or gets widened, these obstacles might be less of a problem. A certain width might also not be as effective in reducing crash risk once the light conditions are worse. These direct and indirect relations will be tested.

There are also interactions between pillars, in the form of different combinations of elements out of the pillars. Between the road user and a vehicle, an elderly person on a regular bicycle or an e-bike may for example not yield the same risk. Someone can also have plenty of experience with a certain vehicle or fail to steer away when needed, or certain vehicles might be more forgiving in the case of a crash. Between the road user and infrastructure, you can think of someone with either good or bad sight: there is a chance the latter overlooks a bollard (which is related to the research by Schepers and den Brinker (2011)). Between infrastructure and a vehicle, some vehicles are not allowed to ride on a road stretch or better able to handle less smooth surfaces.

Next to the pillars, there are external factors that can impact them, and thus the risk of bicycle crashes. Among these factors are the political, economic, social, technical and physical environments, that can lead to rules and regulations (road safety measures) or change aspects of the pillars and/or risks. If a crash happens, this can feed back to the external factors. Things such as climate change can also pressure governments to adapt or revise regulations. People can become more risk-averse when travelling. Surroundings, such as a school or shopping area, and the weather or time of year are examples of the physical environment that might affect risks. In particular, the total duration of daylight will be looked into here. It affects when street lighting is turned on, and makes that obstacles are less visible. Lastly, exposure needs to be taken into account: the more someone exposes themselves to a street and other people (as expressed in the street length and intensities), the higher the risk they will get involved in a crash.

Information about the blue elements of this conceptual model, namely the infrastructure with obstacles, width and street lighting, external factor (natural light), exposure and bicycle crashes needs to be acquired in order to conduct research on them following a chosen methodology. The next chapter will elaborate on that, also slimming down the full conceptual model to a model specification.

3

Methodology

With the first research question about literature answered, it is time to move to the methodology as a preliminary step before modelling takes place and results can be shown. Information about the location and gathering of several data types for that location will be expanded upon, after which the data processing is discussed. The actual modelling approach to make use of the data is then elaborated. It is based on earlier studies in the same field.

3.1. Data collection and assessment

3.1.1. Location

The study focuses on crashes on distributor roads with a speed limit of 50 km/h in Amsterdam. In 2015, the infrastructure of the Centrum, Oost, Zuid, West and Nieuw-West districts were assessed. In 2016, the remaining and more outer districts of Noord, Zuidoost, Westpoort followed (Wijlhuizen et al., 2017). Figure 3.1 shows the location of these areas. These streets are also depicted in Figure 3.2. The blue roads are those that were part of the first group, and the red ones those of the second.

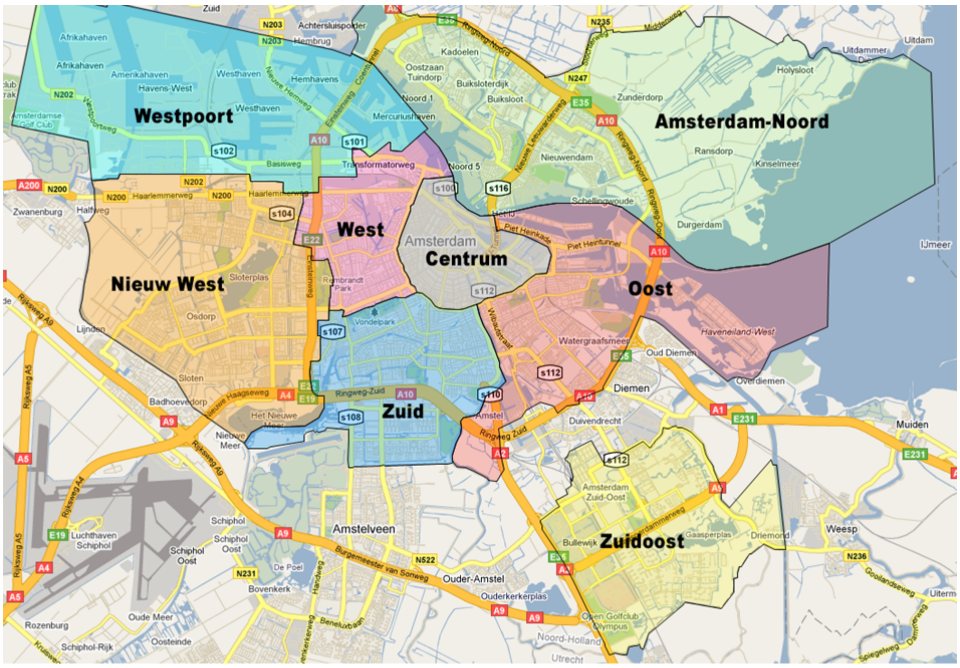


Figure 3.1: Districts of Amsterdam (Klok Real Estate, n.d.)

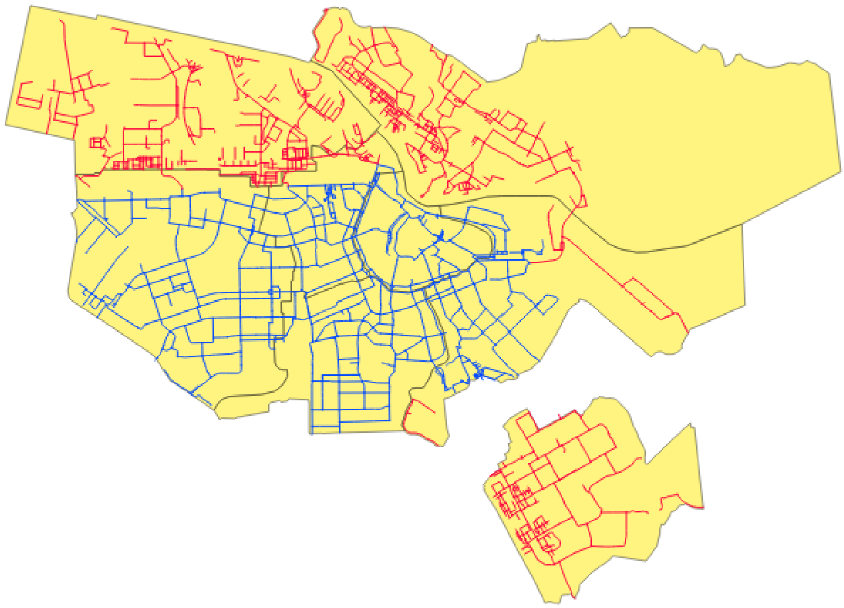


Figure 3.2: 50 km/h roads in Amsterdam (blue depicts the first batch, red the second)

3.1.2. Data sources

To fully assess the effect of infrastructure treatments and thus study the relation between infrastructure and the risk they pose on bicycle crashes, it is important to gather data about the crashes, exposure (such as volumes) and roadway characteristics. For the first, mostly underreporting is a problem (DiGioia et al., 2017). Since SWOV already conducted several studies in which on a large scale data on Amsterdam are obtained and the topic was proposed by them, most data were available in the data set provided at the start of the study.

Crash data

Sources for the crash data are ambulance data. These data are taken from a more raw file and already processed in the data set. Data from the National Institute for Public Health and the Environment (RIVM) from 2009 to 2012 were used in Wijlhuizen et al. (2017). To this, data from the municipality of Amsterdam itself about crashes from 2014 to 2016 are added. As both batches were acquired by SWOV from different sources that provided the data independently from each other, the year 2013 is missing. Nonetheless, 7 years of data are available and it is expected that one year would not significantly change the distribution of crashes over the streets. An alternative, that is more freely accessible than ambulance data, would be the so-called Bestand geRegistreerde Ongevallen Nederland (BRON), that contains all road crashes that were reported to the police and road inspectors of the Directorate-General for Public Works and Water Management (Rijkswaterstaat). This, however, makes that it lacks quite some data where either of those bodies was not involved, especially for single-bicycle crashes that happen to make up most of the bicycle crashes. Together with other organisations, SWOV is working on making these numbers available for traffic safety policies (Rijkswaterstaat, n.d.).

Since not all of these data are logged on exact 4- and 6-digit postal code levels, it is not possible to exactly link certain environments to crashes and the analysis will be done on street level. Given that there is a difference in time periods, it could be that the situation in a street has changed over the years. An ambulance is also not required or called for all crashes, which leads to underreporting. This is especially the case for single-bicycle crashes (Schepers et al., 2017). This is unfortunate but accepted for this research, since most streets are expected to have stayed the same, and crashes without an ambulance are assumed to lead to very minor damage or injuries, if any at all.

Furthermore, when writing about road crashes, an event on a public road is meant, that occurred in traffic and led to damage to objects and/or injury to people, where at least one moving vehicle was involved (SWOV, 2016). The word "accident" is not used, as it is not assumed to be something that happens merely by chance. Since road users such as pedestrians and horse riders do not use a vehicle, by definition these are not categorised as road crashes. For road deaths, the definition of most countries, including the Netherlands, will be applied: people that die within 30 days due to a road crash (World Health Organization, 2009).

Exposure data
Exposure data gives a general idea about how long or with how many people someone can be exposed to the variables to be studied, and that may result in a crash. In this case, these are intensities and street length. The available intensities encompass both motor vehicle intensities and cyclist intensities. The first intensities are taken from the traffic model for 2010 by the municipality of Amsterdam in annual average daily traffic. The bicycle intensities are averages of the intensities gathered during the Fietstelweken ("Bicycle Count Weeks") in September in 2015 and 2016 (Wijlhuizen et al., 2017). More information about the lengths can be found in the infrastructure data.

Infrastructure data
The roadway characteristics were extracted from images taken from CycleRAP and captured by Cyclomedia in 2015 and 2016. These can currently be viewed on the Street Smart website[1]. Only streets that the municipality of Amsterdam classified as a distributor road were considered. Even though a photo was taken every 5 metres, the characteristics were only assessed every 25 metres.

Figure 3.3 gives an overview of how the assessment was executed. 25-metre sections were assessed and put together in the data set. Sections that were fully in an intersection were left out. Next to the instructions for the annotators to assess the included cycling infrastructure and a full list of the infrastructure variables that can be found in Appendix C, further details about the process are present in Wijlhuizen et al. (2017).

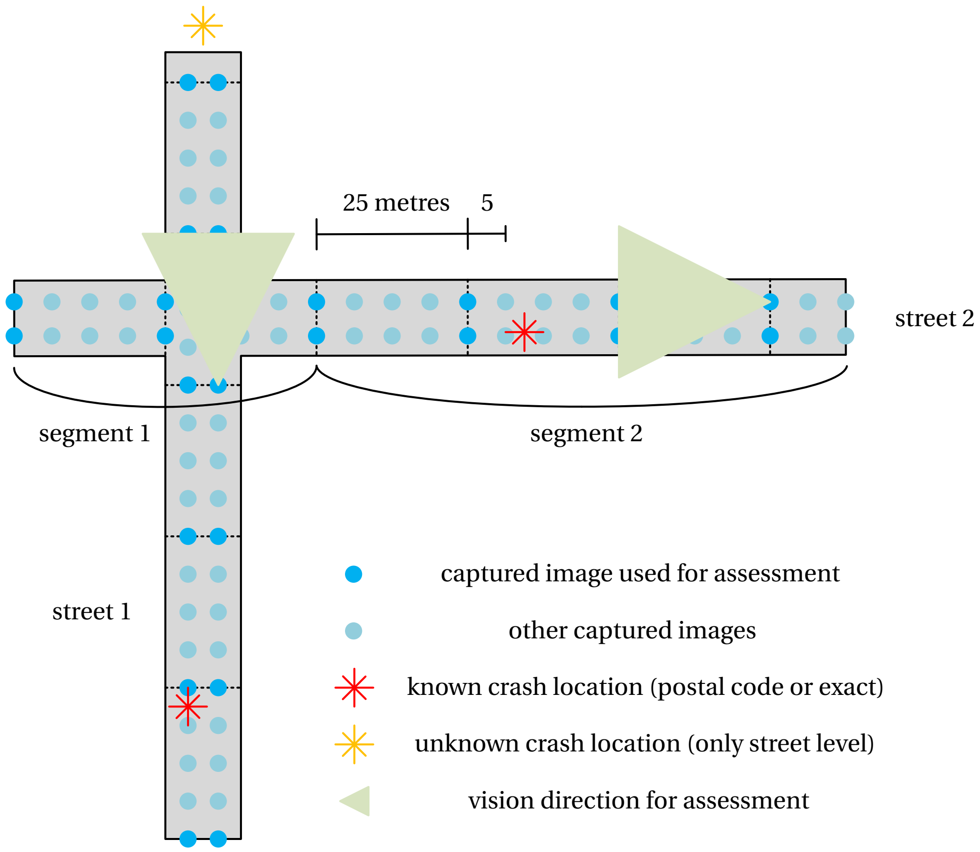


Figure 3.3: Visualisation of assessment of infrastructure characteristics

[1] For more information, check https://www.cyclomedia.com/en/software-and-services/streetsmart

3.2. Data processing

3.2.1. Filtering

Not all data are usable. Therefore, some operations were performed to take out the parts that are. For the best overview and to see how many of the available records are used in the coming analysis, it was decided to start with the infrastructure data set, as that contained the most streets, namely 710. From there on, the number of streets was gradually reduced by performing several steps that are depicted in Table 3.1. To match the infrastructure with the crashes, only the streets that have any reported crashes were kept. Since the focus is on bicycle crashes, this is not enough. A further reduction to only streets with bicycle crash records inferred from the data, including zeros, brings the total down by more than 200 streets. The next step, that was also mentioned in Wijlhuizen et al. (2017), is looking at only streets that have a speed limit of 50 km/h for most of the stretch. Most severe crashes namely happen on these streets in cities (Stipdonk & Bos, 2014). Streets with a speed limit 50 km/h over at least 90% of the sections are selected. By filtering out small streets with a length shorter than 150 metres (an arbitrary number taken from the aforementioned study), several streets with special layouts, like bridges and squares, are eliminated. Since some streets might have missing data that make it impossible to do a full analysis of these streets, this is also something to take into account. While at this stage no additional variables were constructed yet to further filter the data, the table does already show the final set of streets used later. Also keep in mind that software packages, such as SPSS used here, take care of these missing data automatically.

Table 3.1: Data filtering

	Filter	Data set	Do not comply	Entries remaining
0	None	Infrastructure	N/A	710
1a	Have general crash data	Crashes	88	622
1b	Have bicycle crash data		142	480
2	Speed limit at least 90% 50 km/h		93	405
3	Minimum street length of 150 m		62	361
4	No other missing data		41	337

3.2.2. Editing

While the assessment of infrastructure through images was done following certain instructions, it could be that due to human error, observations were mistakenly put in the wrong category in the case of pre-existing options, or manual input was typed incorrectly. Examples are insufficient visibility of medians and bollards that are not even present on a section, or an obstacle-free zone of more than 1 metre, while the shoulder type was labelled as being in the vicinity of the road surface by 1 metre or less. Due to the amount of work to be done, only the incorrect option "Deadly" in the quality of the shoulder for a fence or wall was changed to "Bottleneck". All others are kept as is, since it concerns at most a dozen per category and most are not in the variables that will be used anyway.

In some cases, infrastructure variables could also not be distinguished well enough, leading to a "Could not be determined". This is different from actual missing entries. If desired, the previously mentioned miss-entries can be corrected. These entries, however, solely depends on the collector of the data. Even though images were also captured in later years, again it is not deemed feasible to try and fetch these values, as they amount to at least 3000 fields.

3.3. Modelling approach

To answer the research questions, a method is needed that allows for a way to infer bicycle crash numbers from the elements in the conceptual model, and thus streets and the environment. Statistical models are suitable for this. In statistics, regression analysis namely is used to find relations between a dependent variable and one or multiple independent factors. The use of generalised linear regression models is suggested, taking into account traffic flow and modality (motor vehicle and bicycle). A previous report that worked with the same data set as provided for this study applied this as well. This is a regression model that is more general than the classical regression models and allows other distributions than the normal distribution to be applied to the dependent variable (McCullagh & Nelder, 1989).

In classical linear regression, an independent variable X is used to predict a dependent variable Y, with the following common form: $Y = a + bX$, i.e. X and Y correlate. This regression line tries to follow the data points in a scatter plot (with intercept a and slope b), but this will usually not perfectly fit. Therefore, an error term e is added at individual points, that indicates how far away the predicted and real value are from each other (Bijleveld & Commandeur, 2012). Since we want to know if and how certain cycling infrastructure variables and the external factor light relate to bicycle crashes, the number of crashes is predicted (the variable of interest) by the infrastructure variables (the predictor variables). As such, multiple b parameters with accompanying variables can be added to create a model for multiple regression analysis, extending the previous equation to $Y = a + b_1X_1 + b_2X_2 + \cdots + b_kX_k$.

The number of crashes is not the same on every street; on most streets, none or just a few will happen, and on some multiple. Those will thus be distributed over some function. As a crash either happens or not, they are a whole number (count data) and a discrete distribution is needed. A normal distribution is not discrete, which is why a generalised linear model is chosen. The number of crashes was put in a negative binomial distribution previously (Wijlhuizen et al., 2017), which is a distribution of identical, independent trials with probability p that continues until r successes are reached (Hodges, 1994). In the past, other research has applied this negative binomial distribution for crash modelling as well (Bagui & Mehra, 2019). This distribution handles overdispersion, meaning that the variance is larger than the mean, better than what is called the Poisson distribution (Hutchinson & Holtman, 2005). While the study mentioned in the introduction by Raihan et al. (2019) also used the negative binomial distribution, they applied the zero-inflated version. When collected data can contain a lot of zeros that in the end do not say much about the studied topic (structural zeros: when certain subjects are not at risk for the studied behaviour), this distribution can be used (He, Wang, Chen, & Tang, 2015). Since it is thought that traffic crashes can always happen on streets, this zero-inflated model is disregarded. To check which of the two previous options – a Poisson or negative binomial distribution – fits the sample data best, the goodness of fit of the distributions needs to be determined. For this, the Pearson Chi-Square test can be applied (Pearson, 1900).

Regression weights that do not significantly differ from 0 (proposedly with 95% certainty), and thus do not significantly contribute to predicting Y, can be removed from the equations. To compare models of the same distribution, the Likelihood-Ratio (LR) test and Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) will be used to find the models that represent reality with the included parameters best. With the LR test, the difference in likelihoods is compared at the degrees of freedom of one model that is an extension of another (nested models). AIC and BIC are to be used when the models have different structures (Bijleveld & Commandeur, 2012). AIC can be preferred over BIC when the true model is assumed to be complex and the tested models are oversimplifications (Vrieze, 2012). While this is probably the case – crashes namely do not only occur due to infrastructure, and not just several aspects – both can still be checked to see if they yield the same result. Since the AIC and BIC penalise overfitting (and thus complexity), lower values are better.

The step-wise modelling approach and model formulation of Wijlhuizen et al. (2017) will be used as a starting point for the new research questions, given the similarities in research design. This might make comparing the two studies easier as well. In the following equations, C are the bicycle crashes on a street, b_0 is the intercept, other b's are regression coefficients, L is the length of a street, I are the traffic intensities (mv for motor vehicles and cyc for cyclists) and X's represent other model variables. The length can be seen as an exposure variable, that is generally treated by either putting it in the model as another predictor variable, or by putting it in a (natural) logarithm, with coefficient 1, and adding it as another model variable called the offset. When fitting the model, this makes that the left-hand side is subtracted by the offset. Since both the crashes and street length are natural log-transformed, the properties of this mathematical feature make that crashes are not simply subtracted by the length, but divided by it. This effectively changes the crash counts to a crash density: the number of crashes per kilometre of road length on a street. The predicted outcome then also should be interpreted as the rate of crashes per road length unit (Holsclaw, Hallgren, Steyvers, Smyth, & Atkins, 2015; Yan, Guszcza, Flynn, & Wu, 2009). Since the relationship between intensities and crashes was found to be non-linear (Wood et al., 2013), Wijlhuizen et al. (2017) decided to put the intensities in a natural logarithm as well. This decreases the rate at which crashes occur at increasing intensities ("safety in numbers" as discussed earlier). These measures are believed to be sensible and will be maintained. First, a generalised linear model was fitted on only the exposure variables (the length of streets and intensities), as shown in Figure 3.1. When not taking the infrastructure variables into account, it is namely expected that the number of crashes increases with increasing lengths and traffic intensities.

$$\ln(C) = b_0 + ln(L) + b_1 \ln(I_{mv}) + b_2 \ln(I_{cyc}) + e \tag{3.1}$$

By consecutively adding the independent (infrastructure) parameters separately next to the length and intensities as shown in Figure 3.2, possible model improvements will be checked. This will be done on a per-research question basis. As the variables appear in the world together, a model combining them should be developed as well.

$$\ln(C) = b_0 + \ln(L) + b_1 \ln(I_{mv}) + b_2 \ln(I_{cyc}) + b_3 X_1 + b_4 X_2 + \cdots + b_z X_{z-2} + e \tag{3.2}$$

Next to adding these variables as separate entities, the models will be extended with interaction variables. These are a combination of two variables, that take into account the effect of characteristics that occur together on the outcome (Bijleveld & Commandeur, 2012). Obstacles, infrastructure width and light conditions are hypothesised to have an influence on risk and thus crashes alone, but also when appearing together. For example, a street with many obstacles and a narrow bicycle track could lead to more crashes than a street that has few and a narrow path, or many and a wide path.

Due to the extensive nature of the performed assessment, a lot of variables were recorded. It could be the case that some of these variables are more common to appear next to certain values of other variables, leading to multicollinearity. This means that the variables are correlated and thus overlap with one or more other variables. This influences the regression analysis. The extent to which the predictors show multicollinearity therefore needs to be checked. A measure to do this is by looking at the Variance Inflation Factor (VIF), which can be calculated by taking the inverse of the tolerance: $1/(1 - R_k^2)$. The tolerance amounts to 1 minus the proportion of variance of variable k that is shared with other variables. Although opinions about the most optimal cut-off point for this value often differ, variables are generally regarded severely correlated with a VIF higher than 10 (with a tolerance of 0.1), while values lower than 4 (tolerance 0.25+) and closer to 1 are preferred. These, however, are merely rules of thumb, and should not immediately rule out a model (Bijleveld & Commandeur, 2012; O'Brien, 2007).

When assessing the effect of interactions, ways to alleviate multicollinearity and lower the accompanying VIFs, are mean centring and standardisation, where the means of every independent variable are subtracted from the variables, and also divided by the standard deviation in case of standardisation. For both, this results in means of 0. In case of standardisation, the standard deviation also becomes 1. Since interpretation of the results with mean centring is still possible in the original variables, this method is favoured by the author. This, however, also is a procedure of discussion, and should not simply be always applied. The context is paramount: either a micro view of multicollinearity (on individual variables) or a macro view (the model fit). It is advised to use mean centring for multicollinearity at individual variables and when main effects *with* interaction effects come into play. Little value is to be found on the macro view (Iacobucci, Schneider, Popovich, & Bakamitsos, 2016, 2017). It is therefore deemed a useful measure here. It should be kept in mind that all independent variables should be centred, that the dependent variable should not be centred, that interaction terms should also not be centred and be composed of the same variables that are included in the model (so centred, opposed to non-centred), and that the unstandardised regression coefficients should be interpreted (Dawson, 2014).

In regression, the interpretation of regression coefficients for interaction terms is not as straightforward as for the main effects. The interaction effects are therefore usually plotted at different values. These are low and high values, or simply (between) one standard deviation away from the mean if no special values are present (Dawson, 2014). Other variables, the conditional variables, should stay the same. The choice was made to keep these at the means, as this means no extra contribution needs to be calculated next to the intercept and offset due to the centring of the variables. Equation 3.3 shows the general model to determine the effect of the variables on the outcome, filled in in Equation 3.4. The b's are the regression coefficients for respectively the independent variable (IV), moderator variable (mdr) and interaction variables (int).

$$\ln(DV) = Intercept + Offset + b_{IV} IV + b_{mdr} Moderator + b_{int} IV * Moderator \tag{3.3}$$

$$\ln(C) = b_0 + ln(L) + b_{IV} IV + b_{mdr} Moderator + b_{int} IV * Moderator \tag{3.4}$$

4 Descriptives

After a look at the literature and methodology for this study, it is time to evaluate the data more and prepare it for modelling, following the provided model specification. The crash data, exposure data and infrastructure will be discussed accordingly. The chapter ends with an overview of the variables that will be used in the models.

4.1. Model specification

The conceptual model that was defined in the earlier part of this study is stripped down to a model specification where only the parts that are of interest for the main question and sub-questions are shown. It is namely rarely possible to study the full world or model, attributed to data availability, complexity and other factors. The result is shown in Figure 4.1. The variables will be taken from the data set as best as possible to correspond to the different parts of the model and are supplemented if needed. Obstacles both on and off the path of cyclists will be considered. For widths, the effect might differ depending on facility type and deviation from the mean. These therefore are also considered.

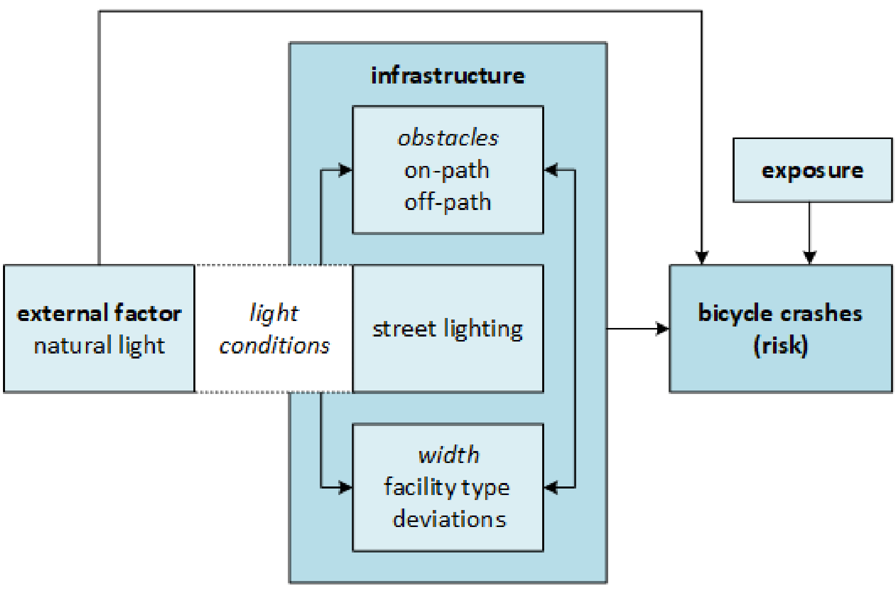


Figure 4.1: Model specification

4.2. Data analysis

4.2.1. Crash data

As mentioned earlier, data about the crashes is taken from a data set previously used by SWOV, that contains both crash data that have been processed from raw ambulance data, and infrastructure data. Since the overarching data set does not contain any information about the time and date of crashes, the raw file needs to be consulted for part of the analysis of the light conditions to be able to take place. The normal data set and raw file are discussed consecutively.

Analysis file

An overview of the crash frequencies of all vehicle types registered by the ambulances in Amsterdam in the filtered set of streets can be found in Figure 4.2. In the case of the second batch, not all streets had reported numbers. For those, simply the numbers of 2009-2012 were adopted. These were mostly streets where no crashes were reported. In general, just over 20% of the streets had no crashes reported.

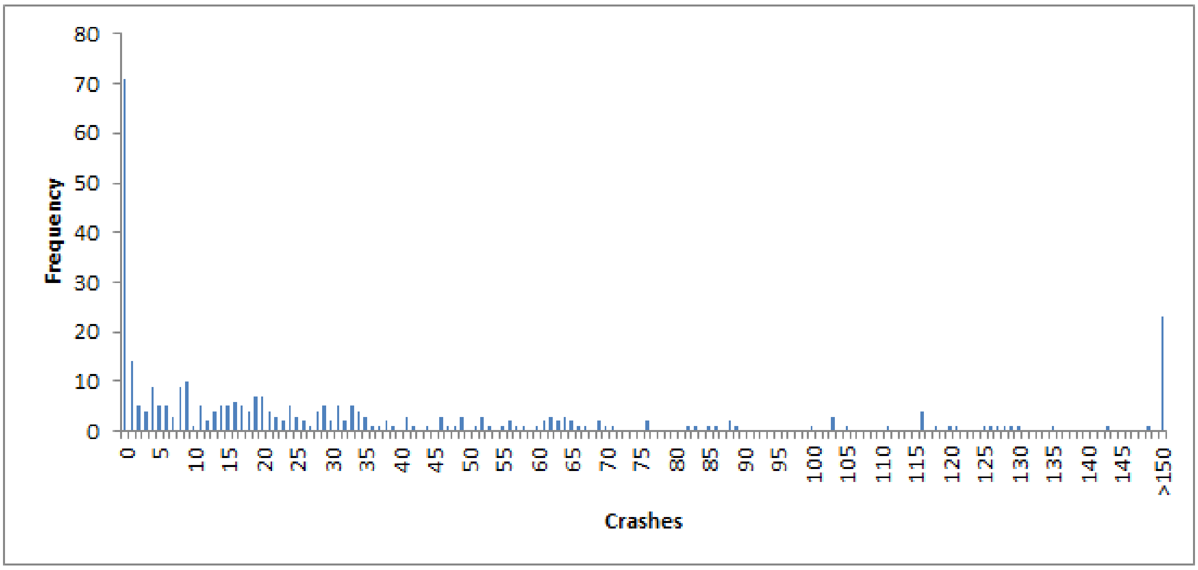


Figure 4.2: Histogram of total crashes on 337 streets from 2009-2012 and 2014-2016

Totals of the crash records after filtering are shown in Table 4.1. Crashes where the mode was registered account for about 42%. Of these, almost 45% of the crashes involved the bicycle. A further breakdown per mode is shown in Table 4.2, with the registered modes also graphically represented in the chart in Figure 4.3.

Table 4.1: Total crash records on 337 streets

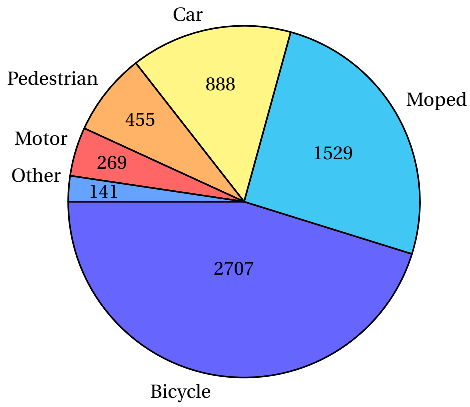
	Ambulance records	Mode records	Cyclist records
2009-2012	7890	4380	2022
2014-2016	-	1609	685
Total	14095	5989	2707

Since the batches have data of different time spans (four years from 2009 to 2012 and three years from 2014 to 2016), it makes sense for the totals of these years to not be in the same range with only minor deviations. Considering both the stagnating traffic deaths numbers and closeness of the years, the ratio of the crash numbers in the two still is expected to be alike the ratio of the duration of the registrations. The provided data only contain numbers for the first batch by itself and the combined batches. Determination of the numbers for 2014 to 2016 did not seem to give reasonable results, as some streets yielded negative numbers. This makes that only the first part can be compared with the total, being about 57% for the year ratio. With a share of 56% for the total crash record, it comes very close. Apart from the sources that provided the data, these similar numbers are a factor that makes it reasonable to assume that the data are reliable. As can be seen in Table 4.1, both the involved mode registration and cyclist crash registration are rather different, however, with the first years accounting for 73% and 75% of the crashes. This could be the result of a difference in registrations of

Table 4.2: Crashes per mode on 337 streets

Type		Crashes	Share
No registered mode		8106	57.51%
Registered mode	Bicycle	2707	19.21%
	Bus	13	0.09%
	Car	888	6.30%
	Light moped	4	0.03%
	Lorry	4	0.03%
	Mobility scooter	61	0.43%
	Moped	1529	10.85%
	Motor	269	1.91%
	Pedestrian	455	3.23%
	Train	2	0.01%
	Tram	54	0.38%
	Van	3	0.02%

Figure 4.3: Shares of crashes with registered modes



(bicycle) crashes, (bicycle) crashes in general or a combination of both. Again though, the data are deemed of sufficient quality, since the unfiltered mode crash numbers of the two batches correlate with a value of 0.878 and the unfiltered cyclist crash numbers of the two batches correlate with a value of 0.863 (significant at the 0.01 level). This means that streets with a higher number of crashes with a registered mode in the first four years also generally show relatively high numbers in the later years and vice versa.

Ambulance file

For the light conditions, the ambulance file is needed as it contains information about time and date. This section will provide an overview of some relevant information. Since it was provided later in the process, some discrepancies in the data were found, that led to the exclusion of crash numbers for the light condition analysis.

In the previous analyses by Wijlhuizen et al., the total crashes from 2009-2012 as found in the raw ambulance data file were used. The only thing that was not properly addressed there, however, is that this data contains duplicate crashes mostly in December of 2011, and one in December of 2012. The total numbers there thus are slightly off. The numbers for 2014-2016 contain duplicates throughout the years (thought to be partly due to other involved crash victims, but also due to errors like in the other period) and are vastly more different from the bigger analysis file. As only the total RIVM values from 2009 to 2012 seem to correspond with earlier analyses (apart from the duplicates in December), the bicycle crashes from this period will be used for the light condition analysis and elaborated here.

The left image in Figure 4.4 shows the distribution of all registered bicycle crashes on the filtered set over the months with the duplicates removed. The winter and summer months show lower numbers than the others, which could be related to fewer people cycling due to work and temperature. From crashes over the weeks (not shown) follows that the crash numbers are fairly even across weekdays. Only a small dip can be found on Sundays.

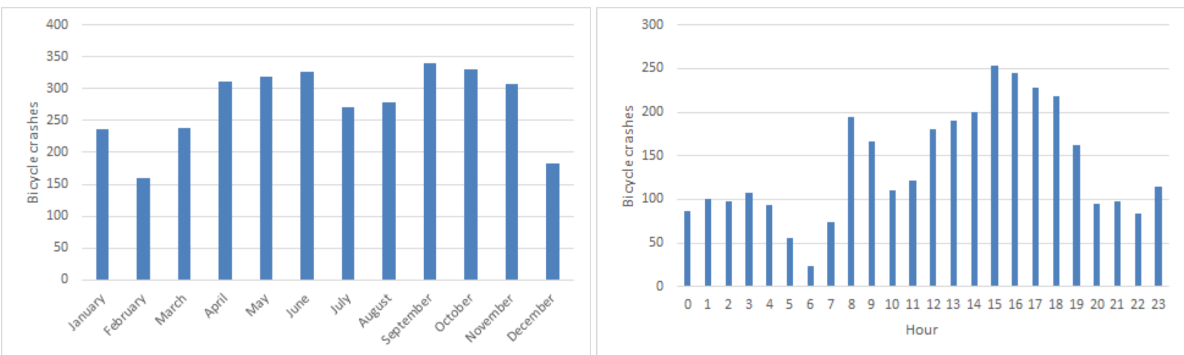


Figure 4.4: Bicycle crashes split per month (left) and per hour on 337 streets from 2009 to 2012

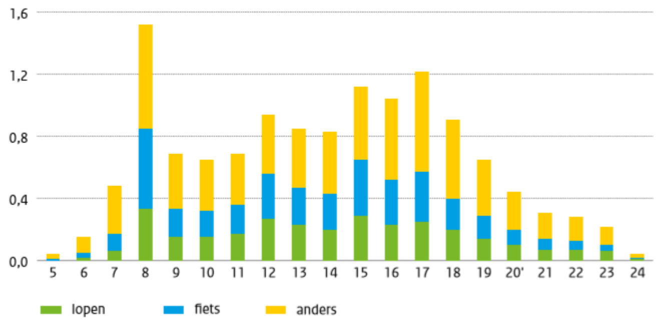


Figure 4.5: Distribution of billion trips over the day for the Netherlands (except weekends), taken from OViN and adapted by Schaap et al. (2015). Green is Walking, blue is Cycling and yellow is Other.

The crashes over the hours in the right image of Figure 4.4 give a fairly logical distribution as well. The nights feature fewer crashes, since fewer people cycle during those times. The highest numbers can be found during peak hours. The morning peak is steeper than the evening peak. This distribution is quite similar to the one found in Figure 4.5, although the morning peak is a bit higher there and the early morning hours are not shown.

Since it is hypothesised to be busiest on the streets during those peak hours, an overview of the crashes according to those is shown as well. The hours used are depicted in Table 4.3. The peak hours are adapted from ANWB (2020) and NS (2020) that are used by the Royal Dutch Touring Club for motorways and the Dutch Railways. Since the evening peak is an hour longer than the morning peak in the first, it was shortened by an hour by taking half an hour from each side. This way they have the same duration. An overview of the crash shares per time of day on the weekdays are given in Figure 4.6. Crashes during the morning peak in the weekend are almost absent, with lots more happening during the night – when people go out.

Table 4.3: Periods split in time of day

	Start	**End**	**Duration**
Morning peak	6:30	9:00	2:30
Rest of day	9:00	16:00	7:00
Evening peak	16:00	18:30	2:30
Night	18:30	6:30	12:00

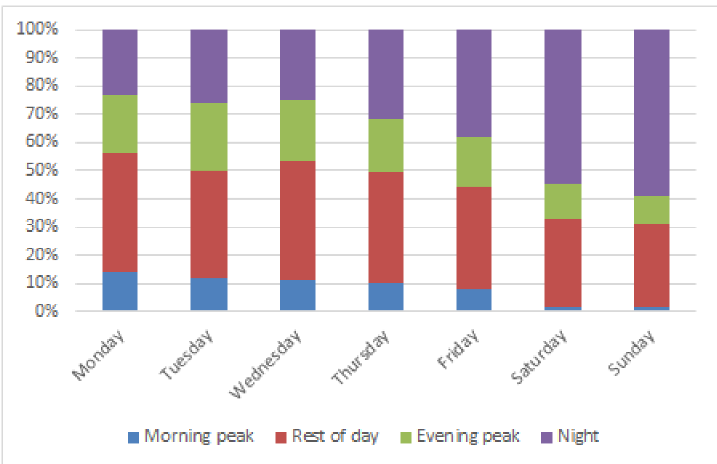


Figure 4.6: Bicycle crash shares per time of day on weekdays on 337 streets from 2009 to 2012

4.2.2. Exposure data

The exposure data are the street lengths and motor vehicle and cyclist intensities, that are added as a correction to the model due to their expected effect, generally independent from the variables that we want to study.

Figure 4.7 shows a histogram of the street lengths of the filtered data. The last bin counts the last few streets that are longer than 3000 metres, going up to just above 5300 metres. The first set of streets between 150 and 200 metres accounts for about 9% of the total streets. This share generally decreases as length goes on, with the mean being around 937 metres.

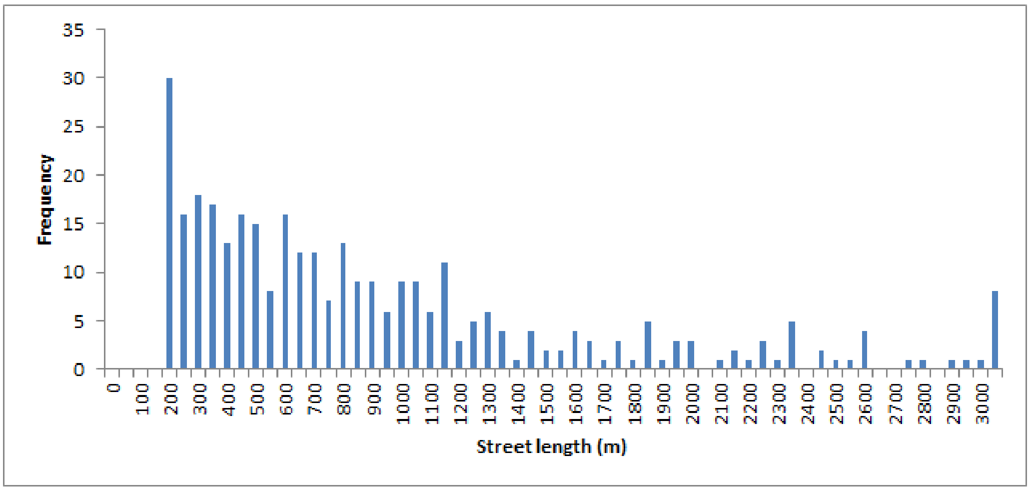


Figure 4.7: Histogram of street lengths of 337 streets from 2015 and 2016

Figure 4.8 then shows a histogram of the motor vehicle and cyclist intensities of the filtered data. The intensities for the motor vehicles are given in hundreds. According to the traffic model, on average most streets (just over one quarter) have to cope with less than 1500 vehicles per day. The minimum is an arbitrary value of 30: it is namely not expected that a street is not used by any cars at all, and the values below 30 were generated by a model anyway. The maximum is 31906 and the mean is an average of 5797 motor vehicles per day. For cyclist intensities, it seems that during the Bicycle Count Weeks, more than 40% of the streets faced 30 or fewer cyclists. Again a minimum was set; this time at a value of 5. The maximum is 635 and the mean is about 87 cyclists in those weeks.

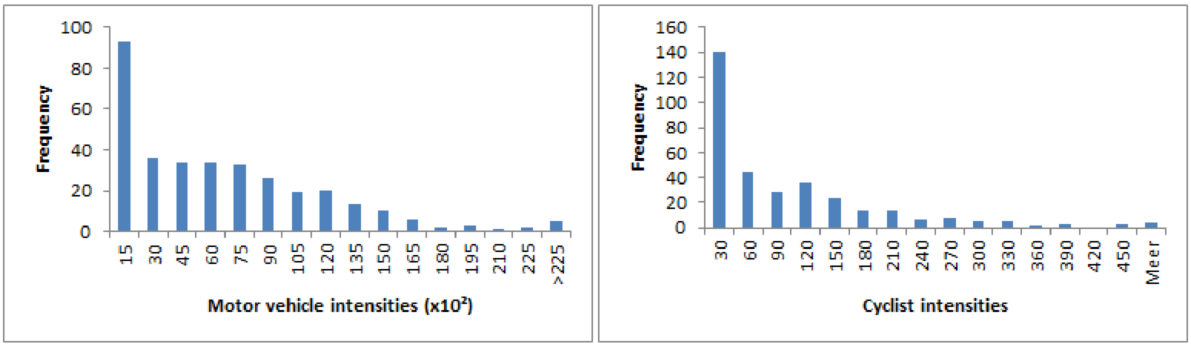


Figure 4.8: Histograms of motor vehicle intensities from 2010 (left) and cyclist intensities from 2015 and 2016 on 337 streets

4.2.3. Infrastructure data

As can be seen from the pie charts in Figures 4.9 to 4.12, the assessed streets in Amsterdam include a broad range of elements. The tables in Appendix C show all characteristics with possible input as present in the data set for both cycling infrastructure and regular motor vehicle infrastructure. For later reference, the last column indicates whether the characteristic was used in the models. The following paragraphs expand on several types.

When looking at the facility types (Figure 4.9), it is clear that two dominate: the biggest shares are taken by the bicycle tracks and carriageway, each taking about 48% and 25% respectively. As follows from the earlier Table 2.2, carriageways are not suggested for cyclists on distributor roads at all. This high share is therefore surprising. The width of the separate types and how these were handled in the model are found in a later section.

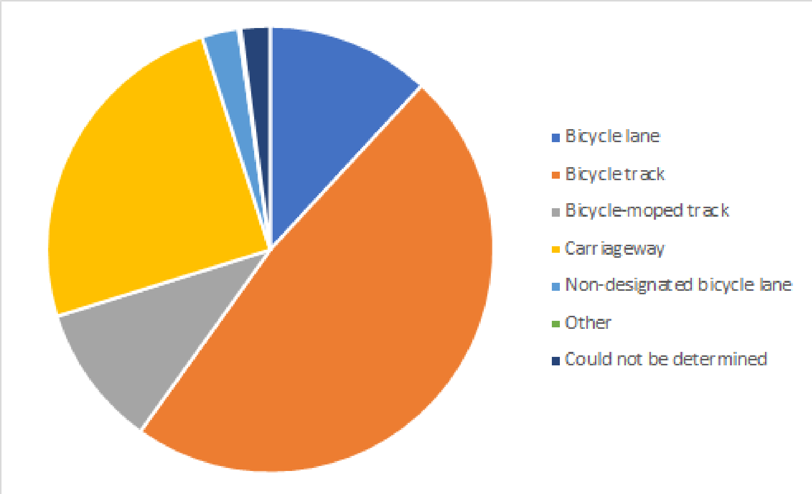


Figure 4.9: Facility types of sections of 337 streets

More than half of the transitions from the road surface to the shoulder as depicted in Figure 4.10 are formed by kerbs, split about equally into impassable and passable kerbs. The next third is split between parking on and along the carriageway, and no real transition at all, namely flat. The quality of these transitions depends mostly on the height difference (also depending on cracks, holes and bumps), and whether there is a chance to lose balance. For bottlenecks, contact should be avoided, while it is uncomfortable for points of attention. 48% was found to be a bottleneck, 80% including points of attention. Given the fact that points of attention are more related to comfort and not direct danger, only bottlenecks will be taken into account for the models. This also makes it easier to discuss implications.

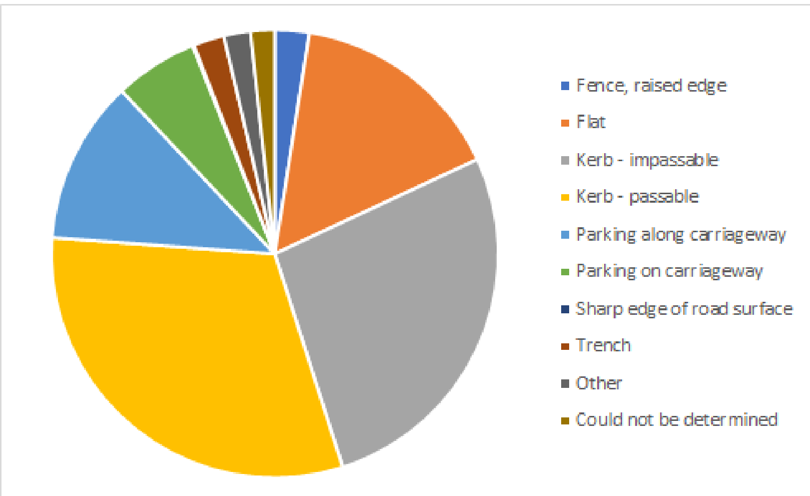


Figure 4.10: Transition types of sections of 337 streets

From Figure 4.11 follows that the most common types of shoulders are parking places, road surface such as pavement, and grass. These take up a quarter and two fifths. While not every type might seem a threat at first, still a significant share of 81% of these shoulders is labelled as a bottleneck, increasing a little to 89% including points of attention. Again, only the bottlenecks will be considered in the models.

Almost 6 out of 10 sections feature obstacles that are 50 centimetres or closer, even adjacent, to the road surface (see Figure 4.12). Furthermore, a meagre 0.6% of the street sections contain a median (but 42 streets, or 12.5%), and nearly 0.1% contain a bollard (but 17 streets, or 5%). Given their low occurrence, this could already be a

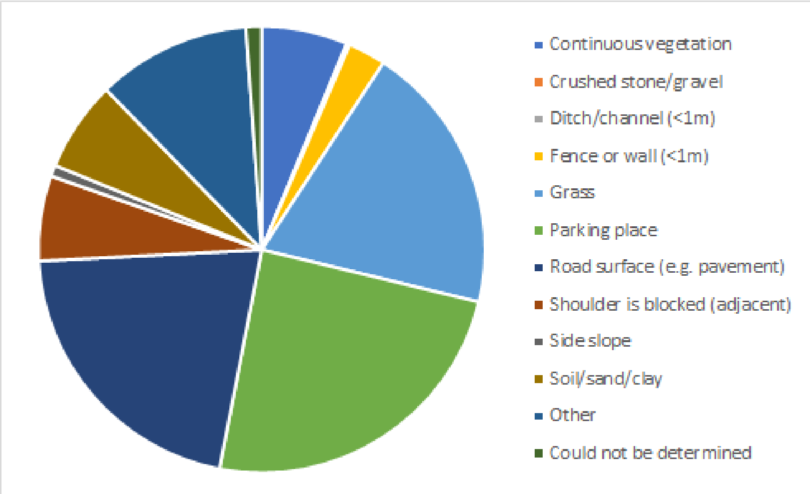


Figure 4.11: Shoulder types of sections of 337 streets

sign that it might be hard to find a relation between these obstacles and bicycle crashes from this data set. Also, around 9% of the street sections do not have lamp posts.

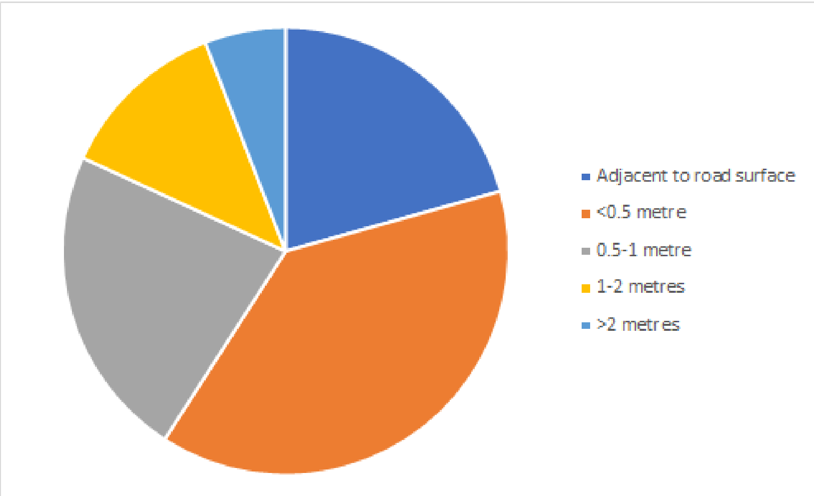


Figure 4.12: Shortest obstacle-free zones of sections of 337 streets

As a check on the validity of the data set available for this analysis, data from the Road Safety Comparison Tool (Verkeersveiligheidsvergelijker) by Fietsersbond, SWOV and VVN (2020) was used. This only concerns bicycle tracks, however. The biggest differences are found in the number of directions on bicycle tracks, the presence of access restrictions for cyclists on carriageways and the pavement quality of bicycle tracks. Apart from those, the other categories are much closer to each other, and thus fairly representative. While not everything can be explained, it should be kept in mind that they also concern 80 km/h roads. A more elaborate discussion on this comparison can be found in Appendix D.

4.3. Variables

4.3.1. Obstacles

Data

In the data set, the obstacles mentioned in Chapter 2 (lamp posts, traffic signs, bollards, fences, walls, kerbs, trees and parked cars) are also present in some way. The previous section already gave some examples. Mainly the categories "Type/Quality of shoulder" (all types except kerb), "Type/Quality of transition" (kerb and more), "Obstacle-free zone next to bicycle facility" (all types), "Bollard on road" (bollard) and "Median on road" (kerb) contain information about these obstacles. There are a lot of different types of obstacles in the shoulder and transition, of which some are more common than others. Because of this, the choice was made to only focus on the overarching quality of these elements for now, as Schepers et al. (2014) reported this to be important as well. Apart from some modifications to make the data suitable for the analysis model, the obstacle part as pictured in the conceptual model should be covered.

To make sure the models contain mostly independent variables, a look at the correlations between variables was taken. For background, Appendix E contains tables with these correlations. Absolute correlations higher than 0.5 are generally considered moderate to high (Hinkle, Wiersma, & Jurs, 2003). For these reasons, it was decided to prevent several combinations from appearing in the models that showed correlations with these and higher values: only shoulder quality when the obstacle-free zone share for the closest obstacles adjacent to the road surface was present, and no quality of the transition in these cases. Naturally, only one type of obstacle-free zone share should be added to the models. Correlations between those thus do not matter. Correlations for intensities and shares of medians and bollards in a street were sufficiently low and therefore kept in any way.

Conditionally relevant variables
Since the visibility of medians and bollards is also captured in the data set, and these variables might be a useful addition to the models in relation to the light conditions, it was decided not to disregard them and check whether they help explain the bicycle crashes in the models with obstacle variables. When these are added, the variables for the presence share of these obstacles are removed, since to some extent they contain the same information. The problem rises, however, that visibility of obstacles is only relevant when these obstacles it says something about are present. When medians or bollards are present on the road, their visibility can namely be either good or bad (as a simplified case in this example), but when they are not, the visibility is undefined. Dziak and Henry (2017) refer to these nested variables as conditionally relevant variables and present a way to include them in a model. By creating a dummy for the presence of the obstacles and including this both as a main effect and as an interaction with the visibility, without a main effect for the visibility, this visibility can be accounted for. This is generally not common, but justified by the study. As seen in Table 4.4, the previously mentioned visibility variable is altered (Visibility*), and instead of undefined is 0 as well when there is no obstacle. This leads to an interaction term that is 0 in case visibility is good or irrelevant, and 1 when bad.

Table 4.4: Coded values for visibility-related groups (median and bollard)

Group	**Presence**	**Visibility**	**Visibility***
Obstacle on road, good visibility	1	0	0
Obstacle on road, bad visibility	1	1	1
Obstacle not on road	0	Undefined	0

Then, as the visibility interactions are to be included for obstacles, a look needs to be taken at the VIFs. The tolerance and VIF can be acquired from SPSS by performing a regular linear regression. A set of variables has been included, of which most will not interact, just to show the overall effect. The results can be seen in Table 4.5. Most values do not change, apart from dummies that are involved in the interactions and the interactions themselves. However, the interaction statistics were not high to begin with. This might be because the main effect for visibility was excluded from the model, as stated before.

4.3.2. Width

Data
The width of the cycling infrastructure is also present in the data set. However, there are some restrictions and things to take into account. While recording the width of the road surface for cyclists at road sections from the Cyclomedia images for the studies by Wijlhuizen et al., the annotators generally measured the smallest width at every 25-metre section and had to note the full width of two-way bicycle tracks and 1.5 metres as widths at carriageway sections (as found in Appendix C). The next section describes the changes that were deemed necessary to cover the width in the most effective way.

Preparation
Determining

Since we want to find out the actual impact of the width of (bicycle) infrastructure on bicycle crashes, the

Table 4.5: Collinearity statistics for obstacle parameters with and without mean centring

	Non-centred		*Centred*	
Model parameter	**Tolerance**	**VIF**	**Tolerance**	**VIF**
Motor vehicle intensity	0.538	1.860	0.538	1.860
Cyclist intensity	0.513	1.951	0.513	1.951
Shoulder bottleneck share	0.579	1.728	0.579	1.728
Transition bottleneck share	0.834	1.200	0.834	1.200
Obstacle-free zone - 50 cm	0.557	1.794	0.557	1.794
Median share	0.475	2.104	0.475	2.104
Bollard share	0.555	1.801	0.555	1.801
Median dummy	0.539	1.855	0.544	1.838
Bollard dummy	0.455	2.199	0.465	2.153
Visibility problem share median * median dummy	0.659	1.517	0.695	1.438
Visibility problem share bollard * bollard dummy	0.634	1.576	0.657	1.522

space cyclists have at their disposal needs to be derived from the values in the data set. As mentioned, widths of two-way tracks were measured for the full width. This was also found to be the case for two-way lanes – although these occurred way less, and mostly at crossings to cover the distance from one track to another. To only look at the actual road surface width cyclists have at their disposal, the widths at sections with those two-way facilities was halved. On the contrary, the available width for cyclists on carriageways was not that easy to derive and had to be calculated from the data, as space remaining after subtracting motor vehicle widths at the available width. This could lead to negative cycling widths. A description of this process is found in Appendix F.

Defining

In order to be able to keep enough streets in the model, and considering the similarities between or differences in several bicycle facility types, it was decided to merge and split some of the types. Non-designated bicycle lanes and bicycle lanes were combined into the simpler "lane" type, bicycle tracks were separated in the "one-way" and "two-way" track types and carriageways were kept as-is. The other bicycle facility types (Could not be determined and Not applicable) were left out.

Since streets furthermore typically do not contain only one type and we want to look at the impact of facility widths without compromising or negating width differences among streets, at the same time it was decided to look at streets with one bicycle facility type that was clearly dominating; that is, more than 50% available along the street sections. This should not lead to too little streets in the model, but at the same time keep the results as pure as possible. To make this possible, dummy variables for each facility type should be created:

- $DUMMY_L$ = 1 for streets with lanes as the dominating type and 0 for others
- $DUMMY_{T1}$ = 1 for streets with one-way tracks as the dominating type and 0 for others
- $DUMMY_{T2}$ = 1 for streets with two-way tracks as the dominating type and 0 for others
- $DUMMY_C$ = 1 for streets with carriageways as the dominating type and 0 for others

This makes that $DUMMY_L + DUMMY_{T1} + DUMMY_{T2} + DUMMY_C = 1$. Due to collinearity, one can be omitted (Dziak & Henry, 2017). When simply taking the means of the shares of the four types, an average street has 13% lanes, 29% one-way tracks, 17% two-way tracks and 41% carriageway. Carriageways thus can be seen as the most common type, that also is not specifically designed for cyclists. Hence, it was chosen to be the reference group. The chosen dominance will be explained later.

Table 4.6 shows how the number of streets with a dominating type change as the dominance threshold reduces, along with how many streets get added from the infrastructure data set to the available set for analysis. Logically, the lower the threshold, the more streets are included, and with the filters from Chapter 3 applied, slightly fewer streets will be usable. At 55%, fewer streets are available compared to when the full data set would have been used with only streets that purely contain one facility type.

Including

While the mean of the width of a dominating type itself might be a good indicator, this mean can be subject to changes in the same street. A measure that takes this into account is the standard deviation. Descriptive

Table 4.6: Remaining streets after shifting dominating type threshold

	Total						*Filtered*					
Share	**Lanes**	**Tracks 1D**	**Tracks 2D**	**Carriage-way**	**Streets**	**Diff.**	**Lanes**	**Tracks 1D**	**Tracks 2D**	**Carriage-way**	**Streets**	**Diff.**
100%	32	80	41	189	342		21	47	17	81	166	
95%	35	93	49	195	372	+30	24	59	21	86	190	+24
90%	38	105	53	200	396	+24	25	70	24	88	207	+17
85%	46	118	60	205	429	+23	33	82	29	92	236	+29
80%	51	125	66	211	453	+24	37	87	32	96	252	+16
75%	53	132	68	216	469	+16	39	94	34	99	266	+14
70%	57	139	69	219	484	+15	41	100	34	100	275	+9
65%	60	148	73	222	503	+19	43	106	36	102	287	+12
60%	63	152	79	229	523	+20	46	108	39	106	299	+12
55%	64	162	85	232	543	+20	47	115	42	107	311	+12

statistics of the standard deviations of the previously defined types are shown in Table 4.7. These are not subject to the dominance threshold, and streets can contain multiple types: the sum is therefore higher than the total street count. The variation of these standard deviations is shown in the histogram of Figure 4.13. The minimum is 0 (which was also set for streets that did have means noted later, but no standard deviation present – the software would have thrown out this entry otherwise due to missing values), and the maximum is found to be over 2 for carriageways. The standard deviation of the standard deviations appears to be higher than the mean for all facility types. The histogram shows that for all types there is quite some variation. Next to the theoretical reasoning, this practical aspect means that the standard deviation of the widths is also added to the width-crash models.

Table 4.7: Descriptive statistics of the standard deviations of the widths of facility types

	Lane	**One-way track**	**Two-way track**	**Carriageway**
Count	108	151	99	143
Minimum	0*			
Maximum	1.06	0.96	0.67	2.31
Mean	0.13	0.14	0.09	0.26
Standard deviation	0.21	0.18	0.11	0.38

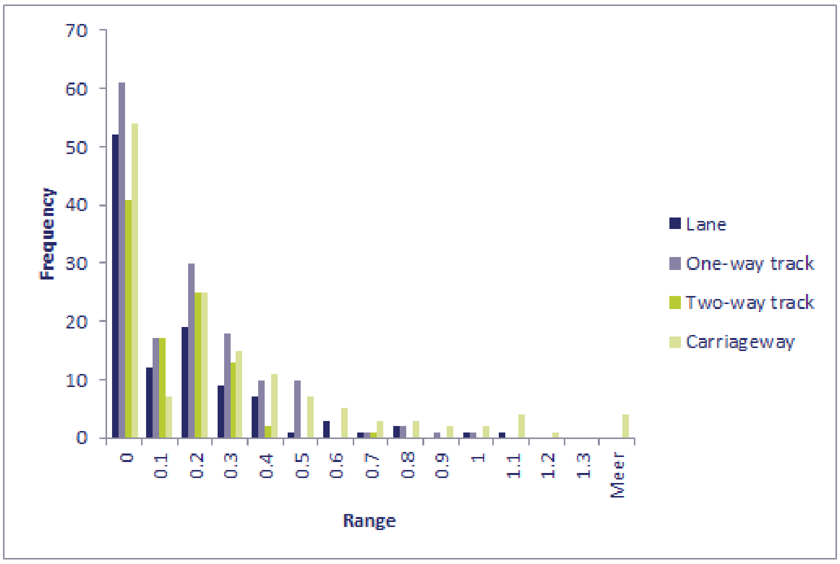


Figure 4.13: Histogram of standard deviations of the standard deviations of widths of facility types

This time, no variables will be left out, as correlations across all combinations were found to be at most slightly moderate, with the highest value being 0.511 between the dummy for one-way tracks and the mean of the widths (not taking the intensities into account). See Appendix E for the full table. Lastly, the width variables need to be checked for collinearity. The dominance threshold was put at 75%, although this did not matter as much for the collinearity indicators. From Table 4.8, it can be seen that both the measures initially are fine for the intensities, mean and standard deviation of the width, and narrowing share, but not for the dummies and interaction variables. These are either very low or very high, and therefore indicate too much multicollinearity. After centring the variables, the variables that were already fine and not involved in the interactions stayed the same. On the other hand, the other variables all drop significantly, indicating that the centring indeed has the intended effect.

Table 4.8: Collinearity statistics for width parameters with and without mean centring (types 75% dominating)

	Non-centred		*Centred*	
Model parameter	**Tolerance**	**VIF**	**Tolerance**	**VIF**
Motor vehicle intensity	0.439	2.276	0.439	2.276
Cyclist intensity	0.414	2.414	0.414	2.414
Dominating width mean	0.406	2.463	0.238	4.210
Dominating width standard deviation	0.693	1.443	0.693	1.443
Lane dummy	0.017	60.532	0.470	2.126
One-way track dummy	0.032	30.863	0.230	4.350
Two-way track dummy	0.034	29.636	0.518	1.932
Dominating width mean * lane dummy	0.017	60.471	0.421	2.377
Dominating width mean * one-way track dummy	0.031	32.280	0.529	1.890
Dominating width mean * two-way track dummy	0.032	31.152	0.642	1.558
Narrowing share	0.944	1.059	0.944	1.059

Combining

Another point of interest is the way the widths were entered in the data set. Since the assessment of cycling and motor vehicle infrastructure was done in both the back and forth directions and added as separate columns in the data set (as left and right), these values had to be treated separately in the analysis of the widths. After determining the means and standard deviations for each side, these were combined per street according to Formulas 4.1 and 4.2 (Higgins, Li, & Deeks, 2019). M are the means, N are the sample sizes and SD are the standard deviations. The subscripts l and r indicate the two groups, combined as c.

$$M_c = \frac{N_l M_l + N_r M_r}{N_l + N_r} \tag{4.1}$$

$$SD_c = \sqrt{\frac{(N_l - 1)SD_l^2 + (N_r - 1)SD_r^2 + \frac{N_l N_r}{N_l + N_r}(M_l^2 + M_r^2 - 2M_l M_r)}{N_l + N_r - 1}} \tag{4.2}$$

In Table 4.9, the descriptive statistics of the widths of the filtered data set are presented. These include the altered carriageway and two-way bicycle facility widths, in both the original types (in italics) and revised types (regular font). The original tracks are not split in one- and two-way tracks, and can therefore not be directly compared with the new types. The means of widths decrease from 2.10 m for one-way tracks, to 1.65 m for two-way tracks and lanes, to 0.93 m for carriageways. When looking at the minima and maxima of the types (and taking the two-way character into account), all ranges of the widths as found in the guidelines can be found. Again, due to taking the available width for cyclists on carriageways, negative widths can be found there: the sections are then merely not suitable for cars and cyclists to drive next to each other.

4.3.3. Light conditions

The natural light that someone is exposed to comes from the sun. The duration this light is available depends on the latitude of the location on earth and the season of the year. The sunlight that is emitted is not only seen between sunrise and sunset, but also in a small time frame before and after that, albeit to a lesser extent. This change from day to night and the other way around is called twilight and can be split into three types, depending on the location of the sun: civil twilight, nautical twilight and astronomical twilight. During these periods, the sun is 0-6, 6-12 and 12-18 degrees below the horizon. In civil twilight, normal outdoor activities can usually be carried out. This becomes harder during nautical and astronomical twilight (Leibowitz, 1987). Data

Table 4.9: Descriptive statistics of bicycle facility type widths

Type	**Mean (m)**	**Number**	**Minimum (m)**	**Maximum (m)**	**Std. deviation**
Carriageway	0.93	3836	-1.15	7.60	0.93
Bicycle lane	1.67	1811	0.75	3.00	0.28
Non-designated bicycle lane	1.54	416	0.60	4.00	0.37
Lane	1.65	2227	0.60	4.00	0.30
Bicycle-moped track	1.81	1628	0.75	3.60	0.45
Bicycle track	2.02	7416	0.90	4.00	0.43
One-way track	2.10	6618	1.00	4.00	0.43
Two-way track	1.65	2426	0.75	3.55	0.27

about the times of sunrise and sunset are taken from KNMI (2019), with the full table shown in Appendix G. Furthermore, data about the duration of civil twilight was found as well (KNMI, 2018). The sunlight conditions that follow from these data are shown in Figure 4.14.

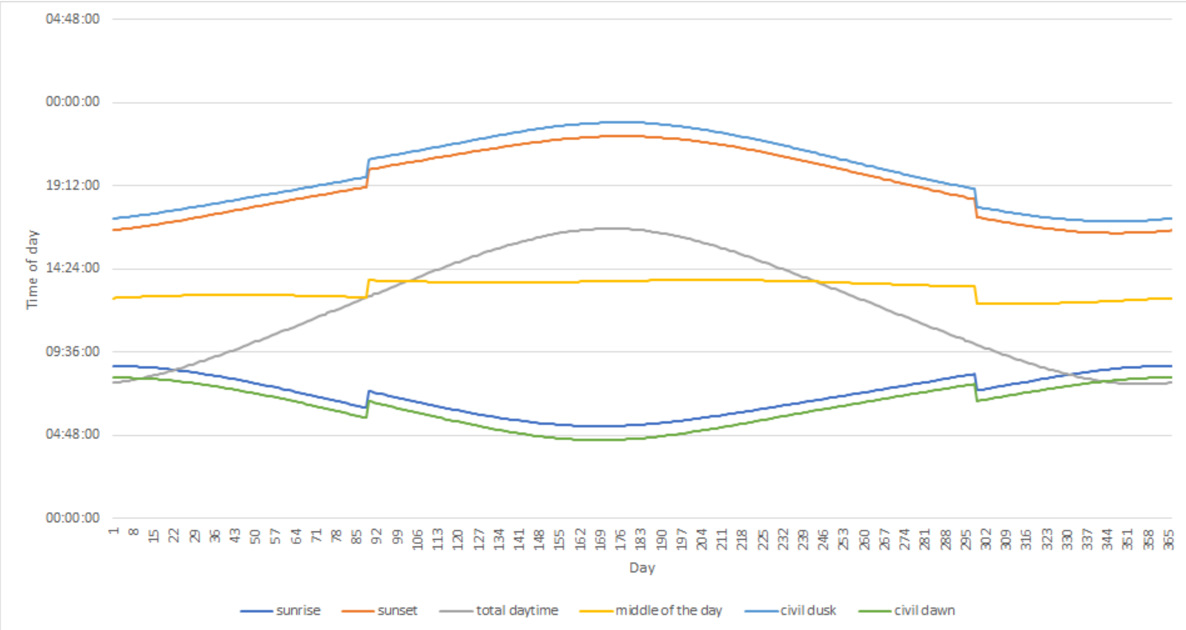


Figure 4.14: Sunlight conditions for 2020 (with data from KNMI (2018, 2019))

There are also human sources of light, such as vehicle lights and lamp posts. These too depend on the location (rural areas might be lit less), and the people in that area. Apart from the lamp posts, these sources are not captured in the infrastructure characteristics or otherwise present in the data set, so there is no way to account for those. The presence of street lights will be used as the only human light variable in the light condition model since this is straightforward and can be captured as a share over the whole street where cycling is possible. Still, when looking at the full infrastructure data set, only 8% of the sections do not have any street lighting present. In advance, the impact of those is therefore hypothesised to be marginal, if present at all.

The natural light conditions are deemed harder to capture. The original plan was to split the model in two by using two bicycle crash groups: one where crashes happened in mostly dark conditions, and one in mostly light conditions. This will correspond to different times in the years. Taking the earlier specified peak hour times into account (6:30-9:00 for the morning peak and 16:00-18:30 for the evening peak) and purely looking at the time of sunset and sunrise, from around March to August these peaks are fully in lit conditions after sunrise or before sunset. These months correspond to the spring and summer periods. It becomes a bit more troublesome to find a time span of the same length for mostly dark peaks. Roughly December and January have the peaks at least three quarters before sunrise and after sunset. Apart from splitting the regression models according to light conditions, the intensity variables should be adjusted by a factor too, as crashes in the September months that were measured do not have the same lighting conditions as other months when people do not start to go to work and school. While relations between the obstacle space and light conditions are assumed and should be studied in light of the gap, later scoping led to a more exploratory analysis with cross tables that only takes into account the general light conditions over the years, where the days are simply longer with regards to sunlight duration in the spring and summer than in the autumn and winter.

4.4. Variable overview

Now that the various elements of the different data sets have been discussed, a quick overview of all variables created and used for the models can be found in Tables 4.10 and 4.11 before moving on to the composed models, the results and interpretation thereof. To make interpretation easier, the variables have not been centred yet. The variables related to the width are given for a more select set of streets given dominance (at 75%) as explained earlier. Compared to Table 4.6, the number of streets for that analysis is off by 1 with 267 instead of 266, due to an inconsistency in capitalisation found for one small street. The effect of this is assumed to be negligible. Scale variables concern interval and ratio levels of measurement. Nominal variables concern dummies.

Table 4.10: Overview of scale (interval/ratio) variables

Focus	Name	Variable	Count	Min.	Max.	Mean	Std. dev.
Objective	Cyclist crashes	C	337	0	105	8.03	14.18
Offset	Road length	L	337	150.11	5331.69	937.49	835.08
Base	Motor vehicle intensity	I_{mv}	337	30	31906	5797.27	5639.23
	Cyclist intensity	I_{cyc}		5	635	86.94	104.31
Obstacles	Shoulder bottleneck share	S_{BN}	337	0.00	1.00	0.83	0.21
	Transition bottleneck share	T_{BN}			1.00	0.57	0.38
	Obstacle-free zone share - adjacent	OFZ_{adj}			1.00	0.25	0.32
	Obstacle-free zone share - 50 cm	OFZ_{050}			1.00	0.63	0.29
	Obstacle-free zone share - 1 m	OFZ_{100}			1.00	0.84	0.20
	Obstacle-free zone share - 2 m	OFZ_{200}			1.00	0.94	0.12
	Median share	med			0.27	0.01	0.03
	Bollard share	bol			0.20	0.00	0.01
	On-path obstacle (OPO) share	OPO			0.14	0.00	0.01
	Visibility problem share - median	VIS_{med}			0.27	0.01	0.03
	Visibility problem share - bollard	VIS_{bol}			0.20	0.00	0.01
	Visibility problem share - OPO	VIS_{OPO}			0.14	0.00	0.01
Width	Dominating width - mean	W_M	267	-1.15	3.68	1.58	0.63
	Dominating width - std. dev.	W_S		0.00	1.90	0.15	0.24
	Narrowing share	nar		0.00	1.00	0.02	0.07
Light	Lamp post share	$lamp$	337	0.00	1.00	0.92	0.16

Table 4.11: Overview of nominal variables

Focus	Name	Variable	Count	Frequency: 0	Frequency: 1
Obstacles	Median dummy	$DUMMY_{med}$	337	295	42
	Bollard dummy	$DUMMY_{bol}$		321	16
	On-path obstacle dummy	$DUMMY_{OPO}$		280	57
Width	Lane dummy	$DUMMY_L$	267	228	39
	One-way track dummy	$DUMMY_{T1}$		173	94
	Two-way track dummy	$DUMMY_{T2}$		232	35

5 Results

This chapter contains the results and implications of several bicycle crash models. Following the research questions and methodology, and the nature of formed variables, every step required running more than one model. Subsequently, models with base variables, obstacle variables, width variables and light variables will be discussed. Combinations are then elaborated. Each section first goes into how the models were formed, how they perform when compared to each other, and what the implications of the results are.

The results are displayed in tables. Green cells indicate significant parameters, while red cells are insignificant parameters with a p-value larger than 0.05 (parameters are found in column B, p-values in column Sig). Dark grey variables refer to parameters that were expected to contribute to bicycle crashes negatively, meaning positive values will increase the numbers. Light grey cells are the other way around. Since interaction effects are not that easy to interpret, they remain white. Significant parameters that clash with the expected direction are coloured yellow.

5.1. Base crash model

5.1.1. Forms

First, several models were run with only the base variables: the offset for the road length (not shown in the results, since these have a fixed regression coefficient of 1 and did not have to be estimated) and the motor vehicle and cyclist intensities. This is not directly formulated in any research questions, but stated in the methodology and used as a first check. The intensities were put in the models in their original form, or in a natural logarithm to simulate the "safety in numbers" principle. A further elaboration on the equations can also be found in Appendix H. The next split was then made in the distribution type, namely Poisson or negative binomial, to see which provided better models. Lastly, the variables were either centred or not to check the influence of them on the regression coefficients and model fit; these should stay the same since there are no interactions present yet. To prevent showing too much data, only the results of models with centred values are presented and discussed further. The results of the base models are displayed in Table 5.1.

5.1.2. Performance

Measures of fit that help indicate the model performance are given below the parameters. It indeed did not matter whether variables were centred or not: both the regression coefficients and measures of fit stayed the same. For consistency's sake and advice in a paper by Dawson (2014), *it was therefore decided to simply centre all variables in future models.* The Pearson Chi-Square value divided by the degrees of freedom is much lower and closer to 1 for the negative binomial distribution (models B2 and B4) than for the Poisson distribution (models B1 and B3). The negative binomial models thus have a better fit to the data. While later models also had one model run with the Poisson distribution, these all gave the same result. To that end, later results do not show these anymore, and only contain models with negative binomial distributions. Since models B2 and B4 have different model forms due to the logarithms, they should be compared through the Akaike and Bayesian Information Criterion (AIC/BIC). The following rule applies for the AIC and BIC: the closer to zero, the better. These are reduced by more than 120 for the negative binomial models that had the intensities in a logarithm,

Table 5.1: Analysis results of base models

	Poisson		Neg. binomial		Poisson		Neg. binomial	
	Model B1		**Model B2**		**Model B3**		**Model B4**	
	PARAMETERS							
	B	Sig	B	Sig	B	Sig	B	Sig
Intercept	-5.09	0.00	-5.14	0.00	-5.67	0.00	-5.52	0.00
Motor vehicle intensity (not in logarithm)	4.2E-5	0.00	6.5E-5	0.00				
Cyclist intensity (not in logarithm)	5.4E-3	0.00	8.4E-3	0.00				
Motor vehicle intensity (in logarithm)					0.14	0.00	0.31	0.00
Cyclist intensity (in logarithm)					0.94	0.00	0.75	0.00
Negative binomial (r)			0.90				0.59	
	MEASURES OF FIT							
Pearson Chi-Square / df	8.219		1.573		6.130		1.383	
Log-Likelihood	-1582.550		-870.826		-1237.430		-809.524	
Akaike Information Criterion (AIC)	3171.093		1749.652		2480.867		1627.048	
Bayesian Information Criterion (BIC)	3182.554		1764.932		2492.327		1642.328	

meaning that the toned-down effect of higher intensities makes the model perform better. *Later models will therefore also have the intensities put in natural logarithms.* Only taking the base variables into account, the best model thus is Model B4. Extensions of this base model are discussed in the following sections.

5.1.3. Implications

In the models, a positive parameter value indicates that a higher variable value leads to higher bicycle crashes on a street, and vice versa. The models to this point do not contain a lot of variables and are still very basic, so interpretation is fairly simple. The regression coefficients in all models remained significant, with a p-value below 0.05. This indicates that streets that process more people, be it in motor vehicles or on bicycles, generally have higher bicycle crash numbers. This makes sense since there will be more opportunities for crashes. It also corresponds to the earlier study by Wijlhuizen et al. with partially the same data. The intensities will be kept in later models.

5.2. Obstacle-crash model

5.2.1. Forms

The first modelling question "What is the relation between obstacles on and around cycling infrastructure and bicycle crashes?" is answered by creating obstacle-crash models, where the base model is supplemented with obstacle variables. The results are displayed in Tables 5.2 and 5.3. Interactions were only added after running models O1 to O5 with just main effects, where the obstacle-free zone was alternated (with the transition bottleneck share only switched for the shoulder bottleneck share for the adjacent obstacle-free zone share) and the combination of medians and bollards in a single variable called on-path obstacles was done later. In models O6 and O7, interaction variables concerning visibility problems, both bottlenecks and points of attention, were added.

5.2.2. Performance

Given that the values of the Pearson Chi-Square divided by the degrees of freedom were much closer to 1 for the negative binomial in each case, as mentioned the Poisson distribution was dropped (and is not shown in the table) after the first comparison in favour of the model with a negative binomial distribution. Due to the different combination restrictions mentioned previously and alternating types of variables, no models are nested forms of another. The Log-Likelihood can thus be disregarded, and the focus here should be on the AIC and BIC. The increase of obstacle-free zone distance to 2 metres resulted in a model with the lowest values. The AIC and BIC are lowered some more by combining medians and bollards to one variable, both in the cases without and with interactions. This is not deemed best, however. The combination namely makes distinguishing between the two obstacles not possible anymore. Furthermore, medians dominate this variable. As mentioned in Chapter 4, there are almost six times more sections that contain medians, and more than twice as many streets. This low occurrence also has not made bollards significant in any of the models, while

Table 5.2: Analysis results of obstacle-crash models

	Model O1		Model O2		Model O3		Model O4	
	PARAMETERS							
	B	Sig	B	Sig	B	Sig	B	Sig
Intercept	-5.54	0.00	-5.55	0.00	-5.54	0.00	-5.56	0.00
Motor vehicle intensity	0.31	0.00	0.33	0.00	0.32	0.00	0.30	0.00
Cyclist intensity	0.74	0.00	0.75	0.00	0.74	0.00	0.75	0.00
Shoulder bottleneck share	0.71	0.02						
Transition bottleneck share			0.21	0.17	0.28	0.06	0.33	0.03
Obstacle-free zone share - adjacent	0.16	0.44						
Obstacle-free zone share - 50 cm			0.64	0.00				
Obstacle-free zone share - 1 m					0.81	0.01		
Obstacle-free zone share - 2 m							2.33	0.00
Median share	3.35	0.06	3.54	0.05	3.17	0.07	2.65	0.12
Bollard share	4.38	0.24	3.31	0.37	4.00	0.29	4.26	0.26
Negative binomial (r)	0.56		0.54		0.55		0.53	
	MEASURES OF FIT							
Pearson Chi-Square / df	1.636		1.503		1.587		1.485	
Log-Likelihood	-803.172		-800.404		-801.330		-797.350	
Akaike Information Criterion (AIC)	1622.345		1616.808		1618.661		1610.699	
Bayesian Information Criterion (BIC)	1652.906		1647.369		1649.222		1641.260	

Table 5.3: Analysis results of obstacle-crash models (continued)

	Model O5		Model O6		Model O7		Model OF	
	PARAMETERS							
	B	Sig	B	Sig	B	Sig	B	Sig
Intercept	-5.56	0.00	-5.59	0.00	-5.59	0.00	-5.55	0.00
Motor vehicle intensity	0.30	0.00	0.30	0.00	0.30	0.00	0.31	0.00
Cyclist intensity	0.75	0.00	0.76	0.00	0.76	0.00	0.75	0.00
Transition bottleneck share	0.33	0.03	0.33	0.03	0.32	0.03	0.35	0.02
Obstacle-free zone share - 2 m	2.32	0.00	2.36	0.00	2.33	0.00	2.44	0.00
On-path obstacle share	5.86	0.07						
Median dummy			-0.15	0.38				
Bollard dummy			0.08	0.79				
On-path obstacles dummy					-0.07	0.62		
Visibility problem median * dummy			4.81	0.06				
Visibility problem bollard * dummy			4.53	0.37				
Visibility problem OPO * dummy					9.71	0.05		
Negative binomial (r)	0.53		0.55		0.53		0.53	
	MEASURES OF FIT							
Pearson Chi-Square / df	1.481		1.481		1.481		1.428	
Log-Likelihood	-797.430		-796.473		-796.897		-799.235	
Akaike Information Criterion (AIC)	1608.860		1612.946		1609.795		1610.470	
Bayesian Information Criterion (BIC)	1635.600		1651.147		1640.356		1633.390	

medians just became significant in one model. It is thus decided that the final model for obstacles is Model 04. Since medians and bollards do not contribute significantly, they were removed from this model, turning into Model OF.

5.2.3. Implications

The first thing to note is that the directions of the parameters seem to stay the same, which is a good sign. Given these similarities, the implications of the results of the best performing model will be elaborated here. Any variables not in this model but present in others will be interpreted too. Like in the base models, higher motor vehicle and cyclist intensities lead to more crashes. A higher share of bottlenecks in both the shoulder and transition from the road surface to the shoulder increases bicycle crash numbers. This makes sense since cyclists will bump into those first if they move too close to the sides or try to move to the shoulders. A higher share of the closest obstacles in an obstacle-free zone adjacent to the road or in a distance of 50 centimetres, 1 metre or 2 metres negatively affects crashes too. Only the share of obstacles adjacent to the road is highly insignificant. Apart from that one, this is expected, as when the closest obstacles in a street are nearer, the probability of crashing into them is also higher. Furthermore, the variables for the share of medians, bollards and a combination of these on-path obstacles in a street are positive, meaning that these also increase bicycle crash numbers, although only the effect of medians really can barely be explained from the data in Model O3, since there is just a lack of these in the observed streets. Then, the dummies for these on-path obstacles are insignificant, and interpreting these does not make sense. They were merely put there for the visibility interactions. These were calculated with the obstacles that were both a point of attention and a bottleneck.

When comparing the size of variable parameters across models, most stay around the same value. The obstacle-free zone parameters, however, change from 0.16 to 2.33 when increasing the distance that is taken into account for these nearest obstacles. For the values from 50 centimetres to 2 metres, this could be accounted to the spread in shares. Due to the cumulative nature of the variable (i.e. the maximum value becomes 1 when taking the nearest obstacles at sections for distances of 2 metres and more), the spread in shares decreases when using further distances as the limit, since the probability that most nearest obstacles in a street are within 2 metres instead of adjacent to the road surface or 50 centimetres increases.

As mentioned in Chapter 3, interaction plots are needed for the interpretation of interaction variables. The best performing model with significant interactions is Model O8 where the medians and bollards were combined. The generic equation adapted to this model, exponentiated due to the Poisson or negative binomial distributions, is shown in Equation 5.1. The variables are filled in, with the crashes C as the dependent variable DV, b_0 as the intercept, the obstacle dummy as the independent variable IV and the visibility problem share as the moderator. As described, the main effect of visibility is not included, since it is a conditionally relevant variable.

$$C = exp(b_0 + \ln(L) + b_{mdr} DUMMY_{OPO} + b_{int} VIS_{OPO} * DUMMY_{OPO}) \tag{5.1}$$

A plot of the equation is shown in Figure 5.1. The visibility problem share varies between about 0 on the left at a quarter of the standard deviation from the mean of visibility problems for medians and bollards together (which is only 0.004) to a standard deviation from this mean to the right, at almost 2% points of attention and bottlenecks at the on-path obstacles. From this image, an interaction effect can indeed be observed. The lines cross and are certainly not parallel. When medians and bollards are present on a street, an increase in the share of problems with visibility also increases the bicycle crash numbers. If there are none, crashes slightly decrease; this is thought to be the case due to rounding errors since no visibility problems can arise. Plots of the insignificant interaction for visibility problems of medians and bollards in Model O6 are added to Appendix I, since the interaction for medians is almost significant and the two can be compared.

5.3. Width-crash model

5.3.1. Forms

To answer the third question about the relation between cycling infrastructure width and bicycle crashes, several models with variables related to the width were run. The results of the models are shown in Table 5.4. The models were expanded in every step. They started in a slim form, with basic width parameters of the mean and standard deviation of the width of the dominating type added to the base model (Model W1), in both the Poisson and negative binomial distribution; again only the latter is shown as it performed better. Then, the model was expanded with the type dummies to first prepare for the interactions (Model W2), then with the interactions between the types and the widths (Model W3), and lastly with the narrowing share that might be another indicator for the width (Model W4).

As a sensitivity analysis, the dominance thresholds were varied from 75% to 95%. This did not appear to have an extreme effect, in the sense that *most* parameters stayed either significant or insignificant throughout the results, with the sign of the significant ones in the same direction. Most changes were found on either end of the range, but mostly considered insignificant parameters again. These are assumed attributable to the number of streets available in the data set. While the parameter values associated with the width (main

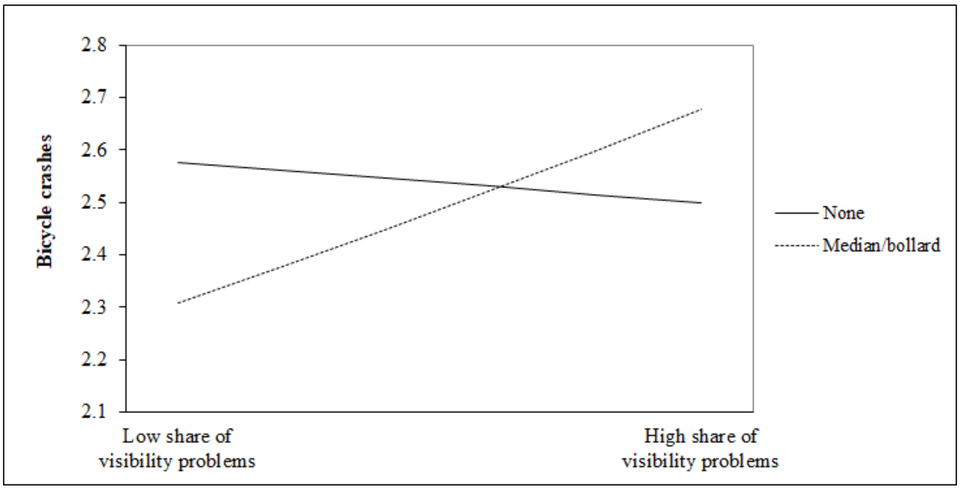


Figure 5.1: Interaction plot for visibility of on-path obstacles

variable and interactions) furthermore stay in the same range as in the models with higher thresholds, the main width variable becomes significant. This is thought valuable for later models that incorporate interactions between width and obstacles. Therefore, 75% dominance was chosen, since the results are based on more data, but the types are still dominating the streets. This way, 70 streets or 20% of the streets available after filtering are removed from the analysis, which is quite a lot. When a lower threshold is chosen, however, that still is higher than 50%, a type simply dominates less and the widths of other types have a higher probability to have attributed to bicycle crashes.

Table 5.4: Analysis results of width models at 75% dominance

	Model W1		**Model W2**		**Model W3**		**Model W4**		**Model WF**	
	PARAMETERS									
	B	Sig	B	Sig	B	Sig	B	Sig	B	Sig
Intercept	-5.62	0.00	-5.67	0.00	-5.50	0.00	-5.49	0.00	-5.50	0.00
Motor vehicle intensity	0.38	0.00	0.34	0.00	0.31	0.00	0.31	0.00	0.31	0.00
Cyclist intensity	0.74	0.00	0.70	0.00	0.71	0.00	0.72	0.00	0.71	0.00
Dominating width mean	-0.20	0.20	-0.17	0.34	-0.42	0.02	-0.41	0.02	-0.43	0.01
Dominating width std. dev.	-0.09	0.67	0.04	0.89	-0.07	0.80	-0.08	0.78		
Lane (L) dummy			0.56	0.02	0.40	0.09	0.38	0.11	0.43	0.03
One-way track (T1) dummy			0.18	0.47	0.30	0.24	0.28	0.28	0.34	0.10
Two-way track (T2) dummy			-0.89	0.00	-1.07	0.00	-1.09	0.00	-1.04	0.00
Dominating width mean * L dummy					-1.66	0.01	-1.66	0.01	-1.66	0.01
Dominating width mean * T1 dummy					-0.99	0.00	-1.01	0.00	-0.99	0.00
Dominating width mean * T2 dummy					-0.31	0.60	-0.34	0.57	-0.27	0.63
Narrowing share							-0.56	0.63		
Negative binomial (r)	0.62		0.43		0.39		0.39		0.39	
	MEASURES OF FIT									
Pearson Chi-Square / df	1.449		1.438		1.345		1.348		1.345	
Log-Likelihood	-616.850		-596.695		-589.183		-589.068		-589.214	
AIC	1245.699		1211.389		1202.367		1204.136		1200.429	
BIC	1267.223		1243.675		1245.414		1250.77		1239.888	

5.3.2. Performance

Again, the Pearson Chi-Square value is much lower for the negative binomial model. The further models and all models shown were thus run with this distribution. Every previous negative binomial model is a nested form of the next, which makes that the Log-Likelihood needs to be assessed. Like with the AIC and BIC goes that values closer to zero are favoured. The Log-Likelihood is highest for Model W4, but the model has three extra variables, so this does not directly make it the best performing model. The Likelihood-Ratio (LR) test, calculated by multiplying the difference in Log-Likelihoods (more extensive minus less extensive) of the two models by 2, has therefore been calculated for all consecutive model pairs: W2 and W1, W3 and W2, and W4 and W3. The results are shown in Table 5.5. The degrees of freedom (df) equals the difference in the number of variables. The critical values are taken from the χ^2 table at a significance level of 95%.

Table 5.5: Comparison of nested models

	LR	**df**	χ^2	**Pass**
W2-W1	40.310	3	7.81	Yes
W3-W2	15.023	3	7.81	Yes
W4-W3	0.231	1	3.84	No

The last model where an extension keeps the LR test above the χ^2 value remains at Model W3. Since the variable for the standard deviation of the width is not significant and not involved in any interactions, it can safely be removed to slightly improve the model. This model is Model WF. Finally, given that the AIC and BIC of the width-crash Model WF (1200/1240) are lower than that of the obstacle-crash Model OF (1610/1633), it can be concluded that the chosen width variables explain the crashes better than the chosen obstacle variables. Still, keep in mind that the obstacle-crash model was run using more data.

5.3.3. Implications

Again, the directions of the parameters stay the same. Given these similarities, the implications of the results of the selected model will be elaborated. Like in the previous models, higher motor vehicle and cyclist intensities still lead to more crashes. A higher mean of the available cyclist width on the dominating facility furthermore lowers crash numbers. This makes sense since a wider stretch to cycle gives more space to avoid potential crashes. Also, note that the width only becomes significant when the interactions come into play. The positive parameter for the lane dummy in the first model it appears in without interactions (Model W2) indicates that streets with these as the dominating types have higher cyclist crash numbers. This was not expected, since it adds a more or less dedicated space for cyclists, that are more designed for cyclists compared to carriageways. The negative two-way track dummy indicates the other way around. Nonetheless, the interpretation of interactions in the later models is still to be done. The effects of the standard deviation of the width, the main effects of one-way tracks as the dominating type, and the narrowing share can unfortunately not be deducted from the data, given their insignificant parameters.

The interaction plots for the width with a facility dummy of Model WF are shown in Figures 5.2, 5.3 and 5.4 to infer the effects of an insignificant and a significant interaction. The equation used for this plot for the streets dominated by either bicycle lanes, one-way tracks or two-way tracks is shown in Equation 5.2, with the dependent variable DV filled in as the crashes C, along with the other variables. The widths in the plots vary from one standard deviation from the mean of 1.58 metres at each side, at 95 centimetres and 2.21 metres.

$$C = exp(b_0 + \ln(L) + b_{IV} W_M + b_{mdr} DUMMY_{type} + b_{int} W_M * DUMMY_{type}) \tag{5.2}$$

An interaction effect for the lanes and one-way tracks indeed is significant. The plots show that a positive change in width on streets with mostly these types has an extra reducing effect on bicycle crashes compared to other streets, especially at lower widths. On the other hand, the last plot for two-way tracks features no crossing lines. While it can be derived that an increase in width reduces bicycle crashes as the main variable for width suggests, the lines are almost parallel. Both reduce at a similar amount at increasing widths. It can thus not be said that an increase in width at two-way tracks specifically compared to other types influences bicycle crashes.

5.4. Light-crash model

Models with the only light condition variable in the infrastructure data set were created: the share of lamp posts. As previously mentioned, the general share in streets is high and it would not be surprising not to find an effect. Table 5.6 shows the only model with the negative binomial distribution and variable added. At a p-value of 0.34, the lamp post share is insignificant. From the data set, it can therefore not be concluded that this variable contributes to bicycle crashes. In the coming combined model, light will thus be disregarded.

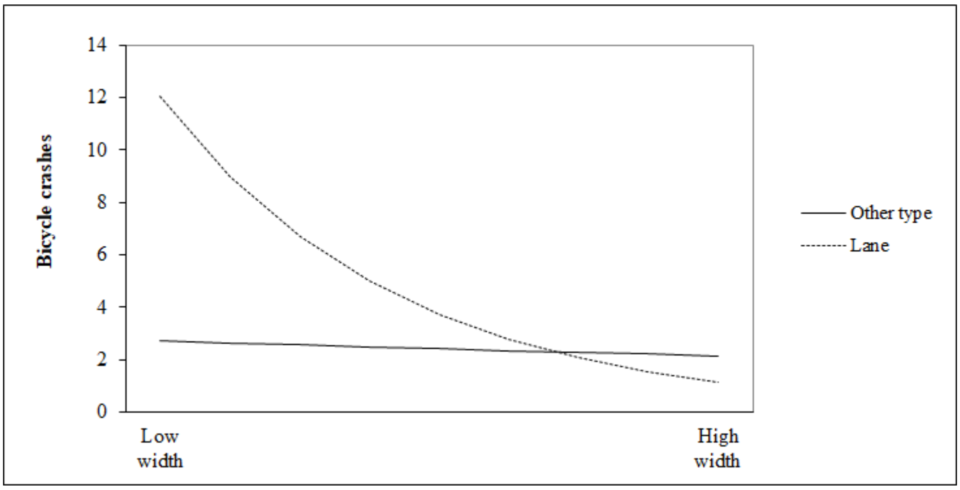


Figure 5.2: Interaction plot for width and lane dummy

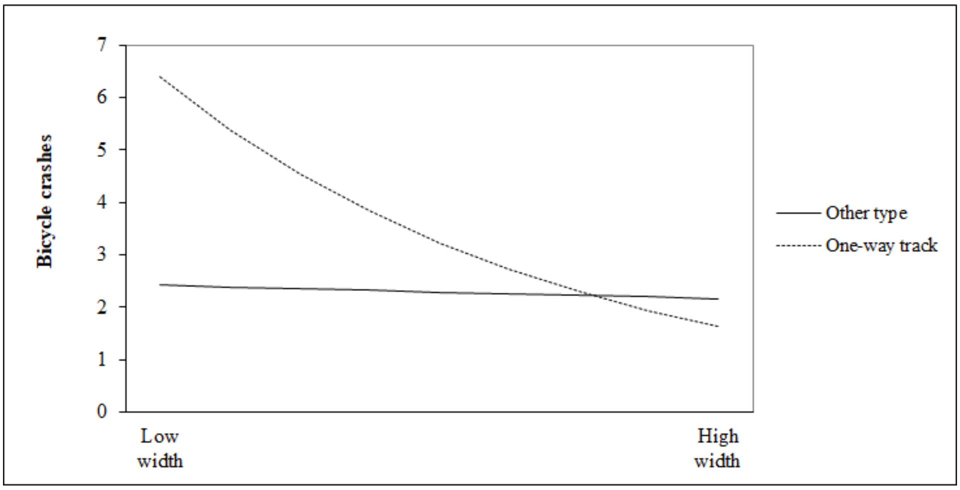


Figure 5.3: Interaction plot for width and one-way track dummy

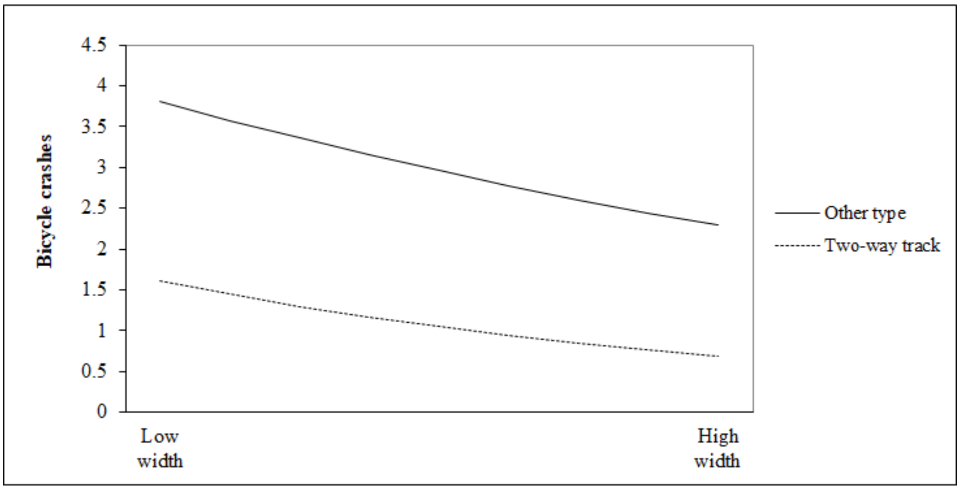


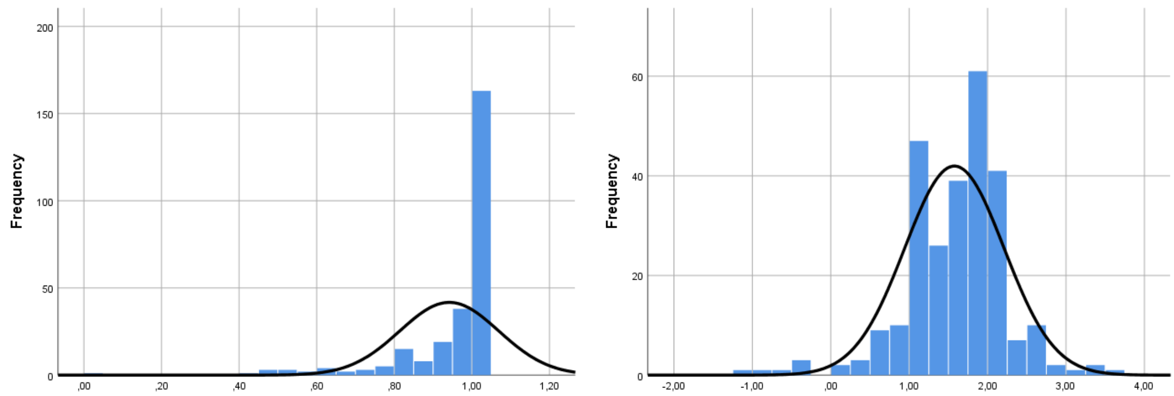
Figure 5.4: Interaction plot for width and two-way track dummy

For a brief look into the different natural light conditions, cross tables are created for bicycle crashes from 2009 to 2012 at streets with different combinations of obstacle-free zones and the mean of the dominating width,

Table 5.6: Analysis results of light-crash models

	Model L	
	PARAMETERS	
	B	Sig
Intercept	-5.52	0.00
Motor vehicle intensity	0.30	0.00
Cyclist intensity	0.75	0.00
Lamp post share	0.36	0.34
Negative binomial (r)	0.59	
	MEASURES OF FIT	
Pearson Chi-Square / df	1.394	
Log-Likelihood	-809.081	
Akaike Information Criterion (AIC)	1628.161	
Bayesian Information Criterion (BIC)	1647.261	

where the date of these crashes is used. As described in the previous chapter, firstly, the total crashes from the later years did not correspond to those in the other analysis files, and secondly, the assumption simply is that crashes that happened in the spring and summer (from March to August) are regarded to have had better light conditions than in the autumn and winter (from September to February). For the categories, depending on the means and distributions as shown in Figure 5.5, set low, average and high values were chosen. The mean of the obstacle-free zone share at 2 metres and width are 0.94 and 1.58 respectively. For the obstacle-free zones, it also is clear that a lot of streets have all of their nearest obstacles in a zone of 2 metres, leading to a share of 1. As low values are not present a lot, the average there was chosen from 0.90 up to 0.95. With 9% and 4.5% of the widths at 1.10 and 2 metres and these values near the 25 and 75 percentile cut-offs at 1.14 and 1.95 metres, the average for the width was chosen between these values from 1.10 metre up to 2 metres.



(a) Obstacle-free zone share at 2 m

(b) Width

Figure 5.5: Histograms of obstacle and width variables at 75% dominance of a facility type

The observed crashes per category are shown in Table 5.7. The combination of a low obstacle-free zone and a low width yielded too low expected values (that were generated by dividing a multiplication of row and column totals by the total crashes in the table), and as such were added to the next category of low obstacle-free zone share and average width. Per the mathematical rules, differences between observed and expected values were squared and divided by the expected values. The resulting values were all summed. This total was compared with the chi-square value in the chi-square table at 95% at the corresponding degrees of freedom ($rows - 1 * columns - 1$, in this case 7). The resulting total of 6.55 is lower than 14.07, meaning that the data cannot indicate a significant difference for bicycle crash numbers in summer and winter, taking the obstacles and width into account. Another cross table was constructed with fewer categories, but that did not point at a significant difference either, even at a lower significance threshold of 90%.

Table 5.7: Cross table of observed crashes for summer and winter on 267 streets

Obstacle-free zone (2 m)	**Width**	**Summer**	**Winter**
Low	Average	47	44
Low	High	50	56
Average	Low	11	7
Average	Average	116	88
Average	High	13	6
High	Low	9	12
High	Average	667	617
High	High	387	371

5.5. Interactions between categories

5.5.1. Forms

The last modelling question concerns combinations of the previous crash models with obstacle, width and light variables, including interaction terms. To get to these models, the final models were simply combined. The dominance threshold of the width-crash models makes that the combined models also use fewer streets. The light-crash models currently did not provide any significant variable and are excluded. The results are shown in Table 5.8. Model C1 features the plain combination. Since the transition bottleneck share, obstacle-free zone share and dominating width mean are the significant main effects in the separate models, interaction effects between these were both added (width * transition and width * obstacle-free zone): first with the obstacle-free zone, since its main effect was still significant (Model C2), and later together with the transition (Model C3) to test if that would change anything. The different levels of the obstacle-free zone variables and switch to shoulder bottleneck share from transition bottleneck share at the adjacent obstacle-free zone allowed for more models, but an obstacle-free zone share at 2 metres proved to perform best.

Table 5.8: Analysis results of combined crash models at 75% dominance

	Model C1		**Model C2**		**Model C3**		**Model CF**	
	PARAMETERS							
	B	Sig	B	Sig	B	Sig	B	Sig
Intercept	-5.53	0.00	-5.54	0.00	-5.57	0.00	-5.53	0.00
Motor vehicle intensity	0.31	0.00	0.31	0.00	0.32	0.00	0.31	0.00
Cyclist intensity	0.71	0.00	0.70	0.00	0.69	0.00	0.71	0.00
Transition bottleneck share	0.00	0.99	0.00	0.98	-0.16	0.51		
Obstacle-free zone (OFZ) - 2 m	1.75	0.00	2.17	0.00	2.46	0.00	1.75	0.00
Dominating width mean	-0.36	0.03	-0.40	0.01	-0.40	0.01	-0.36	0.02
Lane (L) dummy	0.40	0.04	0.43	0.03	0.46	0.02	0.39	0.04
One-way track (T1) dummy	0.22	0.40	0.26	0.33	0.18	0.51	0.22	0.29
Two-way track (T2) dummy	-1.07	0.00	-1.03	0.00	-1.10	0.00	-1.07	0.00
Dominating width mean * L dummy	-1.73	0.01	-1.64	0.01	-1.57	0.01	-1.73	0.00
Dominating width mean * T1 dummy	-0.91	0.00	-0.83	0.01	-0.39	0.37	-0.91	0.00
Dominating width mean * T2 dummy	0.19	0.75	0.04	0.95	0.39	0.54	0.19	0.75
Dominating width mean * transition					0.66	0.16		
Dominating width mean * OFZ 2 m			-1.39	0.12	-1.99	0.04		
Negative binomial (r)	0.38		0.36		0.36		0.38	
	MEASURES OF FIT							
Pearson Chi-Square / df	1.289		1.398		1.347		1.284	
Log-Likelihood	-584.730		-583.436		-582.468		-584.730	
Akaike Information Criterion (AIC)	1195.460		1194.872		1194.937		1193.460	
Bayesian Information Criterion (BIC)	1242.094		1245.093		1248.745		1236.507	

5.5.2. Performance

Just like with the width-crash models, these combined models are nested. The LR test is thus used again. The results of this measures of fit comparison are shown in Table 5.9. Models C1 to C3 were compared with one another, and neither Model C2 nor C3 are found to be an improvement regarding model performance over the previous ones. Model C1 therefore performs best. It is useful to also compare the obstacle- and width-crash models with this model. These are also nested in Model C1. As can be seen from the table, the combined model outperforms both separate models. Removing the insignificant transition bottleneck share from Model C1 subsequently turns it into Model CF.

Table 5.9: Comparison of nested models – combined

	LR	df	χ^2	Pass
C2-C1	2.589	1	3.84	No
C3-C1	4.524	2	5.99	No
C3-C2	1.935	1	3.84	No
C1-OF	429.01	7	16.92	Yes
C1-WF	8.97	2	5.99	Yes

5.5.3. Implications

Since the implications of the separate models were discussed previously, only the changes are elaborated. The most notable differences are the insignificance of the transition bottleneck share and significance of (extra) interaction effects. By combining the predictor variables for different width and obstacle elements, the transition bottleneck share from the data does still not explain bicycle crashes in these combined models. The interaction between the width and obstacle-free zone at 2 metres became significant in one model as well. The other main effects stay the same.

When looking at interaction effects, several things can be noted. The interaction between the width and the type dummies stayed about the same. Since the plots did not change drastically from the previous ones, they are put in Appendix I. From those, again can be concluded that increasing widths of all types decrease bicycle crashes, with the effect having more impact for increases of lower widths. Given the fact that the interaction between width and transition bottleneck share is not significant either, for illustration purposes that plot is found in the appendix too. As mentioned, a significant effect of the obstacle-free zone with the width was found in one model. A plot of that effect is shown in Figure 5.6, that is created in a similar way as Equation 5.2. To stick to reasonable values, instead of going up a standard deviation on the horizontal axis at the obstacle-free zone shares at 2 metres, it was stopped at 0.5 standard deviations to the right for 2 metres, which is about 1. The widths are average (1.58 metres) and one standard deviation from this value (to 94 centimetres and 2.21 metres). Generally speaking, the lines for this interaction indicate that a higher share of nearest obstacles in a 2 metre zone from the road surface leads to more crashes. This effect is larger at lower widths, indicating that more opportunities for swerving and avoiding obstacles at wider sections indeed reduce crash risk.

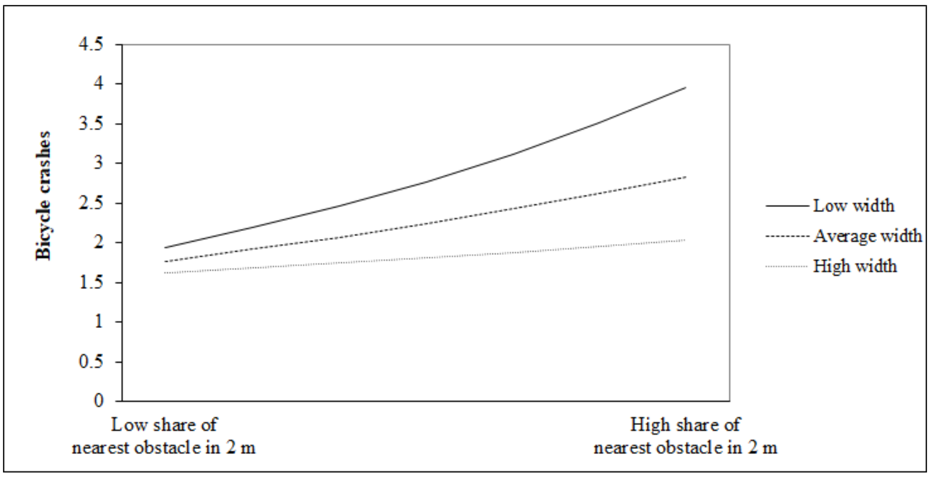


Figure 5.6: Interaction plot for width and obstacle-free zone (2 metres)

6

Conclusions, discussion and recommendations

6.1. Conclusions

While the Netherlands has a low fatality rate for cyclists, the numbers of severely injured cyclists and cyclist deaths were found to be increasing, cyclists make out a big part of the total injured and deaths, and a reduction in total traffic deaths came to a halt. Together with a research gap in the direction of the obstacle space, that relates to obstacles on the road surface and next to bicycle facilities, and the space (i.e. width of the road) that is available to cycle and in the worst case avoid these obstacles, the importance of cycling in the Dutch culture make it a good research topic. This study therefore looked for an answer to the question *"To what extent can the obstacle space of cycling infrastructure and visibility thereof explain bicycle crashes in Amsterdam?"*. For this, bicycle crashes registered by ambulances and infrastructure data extracted from photos are used and changed to street level.

Literature research was used for the first sub-question to gather general characteristics of the obstacle space: the types of obstacles, ranges of widths and determining factors for implementation and crash risk. It became clear that there are many different obstacle types with different functions (such as defining a barrier, controlling access and guarding) and characteristics (on or off the road, fixed or dynamic, and point or line). Crash risk is mainly related to collision, distraction, blocking vision and collision as a result of the latter two. Although data about these types of obstacles was present to some extent, the different nature of all the obstacles for now asked for a more practical approach, only taking the share of nearest obstacles in a certain distance from the road, and the general bottleneck share of the obstacles next to the road into account. Obstacles on the road surface, namely medians and bollards, were defined better and used as a separate entry in the models.

Most common width values from guidelines vary between at least 1 metre for carriageways to 5 metres for two-directional busy tracks. The width of bicycle facilities in these guidelines is dependent on the free-space profile around cyclists and other road users, use of the road regarding directions and overtaking, and intensities. The width of the free-space profile takes into account the maximum bicycle width, sway and fear of obstacles and other road users. It has to be noted, however, that not always these width values are adhered to, for example due to cost or space restrictions. The available width determines swerving possibilities and, again, the risk to collide with obstacles and other road users.

To fill the gap about the relation between obstacles, width, light conditions and the interaction between them with bicycle crashes, modelling of generalised linear regression with a negative binomial distribution was applied. Models with a single focus and a combination of different variables were run. Findings from earlier studies were confirmed: Wijlhuizen et al., that partly used the same data set also found that busier streets lead to more crashes and could not verify the impact of medians and bollards on crashes from the data. Streets that are predominantly two-way tracks again lead to fewer crashes, like in DiGioia et al. (2017); Reynolds et al. (2009).

For the second sub-question, from crash models that used variables only related to obstacles, it was found that more problems in the visibility of medians and bollards increase bicycle crashes, however. The higher the share of nearest obstacles to the road surface in a certain space, the higher the bicycle crashes as well. Also, the more bottlenecks in the transition from the road surface to shoulder, the more bicycle crashes happen. Models with variables related to widths for the third sub-question gave the following results: wider cycling infrastructure has a positive effect on bicycle crashes. Compared to carriageways, the width of lanes and one-way bicycle tracks even has an additional effect on bicycle crashes. No effect of the standard deviation of these widths and share of narrowings in a street were found. All in all, with these results, the hypotheses that more obstacles and narrower infrastructure lead to more crashes were validated. The fourth sub-question gave no indication of a relation between light conditions and bicycle crashes from the data set.

The main hypothesis that more obstacles, smaller widths and less light combined are bad could be verified from a merged model for the fifth sub-question to some extent as well. A changing effect between obstacles and crashes was found for an obstacle-free zone at a maximum of 2 metres away. The smaller the available space for cyclists is, the higher the crash risk with the same share of nearest obstacles in the shoulder. At different widths, no interaction effect was found for bottlenecks in the transition from the road surface to shoulder or shoulder itself. The main effect for bottlenecks in transitions was lost there too.

All in all, clear leads have been found that indicate that to a big extent relations between the obstacle space of cycling infrastructure in Amsterdam and registered bicycle crashes in the city exist on a street level. While not all variables that were included in the models came out significant, many still did. Interaction effects between different elements of the obstacle space were found as well. Apart from visibility problems of on-path obstacles, changing light conditions could not give a reason for crashes in the data set.

6.2. Discussion

First and foremost, a reassuring finding while looking at the results of this study was that the significant parameters for variables were found to be in the expected direction, thus not giving strange results. Despite this, this study can be improved, especially concerning the light conditions. These were now omitted in the combined model, and it is assumed that the effect of obstacles and width is higher in bad light conditions. Although no effects on crash risk could be found for now (resulting in no different effect for light conditions at different obstacle or width levels as well), better light conditions did increase safety in earlier studies. The method as described earlier can be followed or another approach can be chosen.

The conceptual model that formed a basis for the analysis method is bigger than solely considering infrastructure characteristics. The effects of the road user and vehicles, that are also part of it, were excluded fully in this research. This is due to both the focus on infrastructure and the availability of data. Apart from the more general date and time of crashes and in some cases the vehicle type, namely no detailed data about the road user and their vehicle was available, such as age and type of bicycle (especially since older people cycle more during daytime and due to the increase of electric bicycles in the Netherlands). It is useful to control for these in later studies. For example, the use of vehicle lights in dark conditions can be a good addition. Different relations between infrastructure and bicycle crashes can be found, depending on the road user and vehicle type, which helps to determine which adaptations to the infrastructure should receive priority.

Related to the above are the registrations. As can be derived from the data set, with 42 per cent of the registered vehicles and 45 per cent of those being cyclists, a big share of the crashes had to be disregarded. Even though registration can be time-consuming at the crash site, underway to the hospital or after treating a crash victim when it is of utmost importance to decrease traffic deaths and injuries, more detailed information is essential when wanting to make more in-depth studies. Parties that compile this data, such as the ambulance and municipalities, should also focus on making the data more complete.

The location and data sources also determine the applicability of the results. This study specifically concerned cycling infrastructure along 50 km/h road sections (excluding intersections) inside built-up, metropolitan areas of the municipality of Amsterdam. While the initial set of roads are scattered all through the city, most importantly these do not represent all roads where crashes happened, both in the city itself and the country. While some crashes had more detailed locations available that also allow for more detailed analysis of the surroundings that might have led to a crash, all crashes had to be scaled up to the full street level. Albeit rather long, shorter streets in the analyses – 150 metres being the shortest due to filters – result in created variables that are bound to more accurately represent the infrastructure characteristics that were present at the location of the crash. While intersections are a more complex category, ways to analyse those could be pursued too.

A further point to make relates to the quality of the data. The intensities are either taken from an older model or volunteer counts. While the Bicycle Counting Week numbers can actually better correspond with reality, they certainly do not capture full day averages and all streets. Both intensities on streets are thus merely a proxy of the actual intensity that should be used. Nevertheless, these proxies both came out significant and were thus able to predict crashes well. The same can be said for the infrastructure data: road sections can have changed in the meantime since assessing all characteristics. This was already found to be the case in the process by comparing the actual state of several streets with pictures from 2019 with those used in 2015 and 2016. Despite this, crash data and infrastructure are compared around the same years, and any changes that have happened afterwards are neglected.

Just like the crash data that had to be predicted at street level due to missing exact locations, infrastructure data were scaled up to higher levels in the form of proportions and averages as well. Details on every 25-metre section are most likely disregarded, with only the general impression of that section and eventually the streets included. However, it could very much be that one of these details is what led to a crash. For the obstacle-free zones, the closest obstacle on the side of the road was registered. This also thus does not allow for determining the clustering of obstacles on a section, and also does not tell how many are present on that section. For medians and bollards, pretty much the same can be said. Lastly, the different types of transition and shoulders, and thus obstacles, are not fully explored here yet either. All categories in these would also have split the data too much, leading to few cases. The reduction of usable data already became apparent during the analysis. Combining types of the same form could then be an option.

6.3. Recommendations

With recommendations for further research already mentioned in the previous discussion, this section will finally list some recommendations for parties that are responsible for changing infrastructure, such as the municipality. With that, the last sub-question about what measures can be taken to improve cycling infrastructure in the Netherlands and reduce crash risk is answered.

To reduce crash risk, the most straight-forward takeaways from this study are of course that cycling infrastructure should come with as little obstacles as possible and be as wide as possible. Medians and bollards on the road that were points of attention or bottlenecks regarding visibility lacked profiled road markings and general contrast with the environment. For those obstacles, it is thus advised to take the guidelines by the CROW into account.

The quality of transitions (and shoulders) as assessed depended mainly on the height difference between the road surface and the shoulder. Cracks, bumps and holes are what create these height differences. When these create a high probability of cyclists losing balance and contact should be avoided, they were marked a bottleneck. When a transition simply does not exist, which is the case with for example bushes, fences or a quay, there is no space to swerve. The transition in these cases was marked a bottleneck too. These bottlenecks should be prevented and repaired. Maintenance is therefore of great essence.

Instead of making cyclists use the carriageway, it is also better to make any facility like lanes or tracks available. In the case of lanes, this makes sure at least some dedicated space is available, and for tracks this reduces the exposure to motor vehicles that have to use these facilities. Also, this corresponds to one of the measures proposed in the Road Safety Strategic Plan 2030 (Kennisnetwerk SPV, 2019). Several carriageways in the assessed streets even did not have enough width available for all road users to drive next to each other. On streets with speed limits at 50 km/h, speed differences are too high and risk increases. These carriageways also still are frequent on streets. From the interaction plots for lanes and one-way tracks even followed that an increase in average width on smaller sections has a bigger impact on crashes than the same increase on wider parts. It is thus a good idea to take a look at improving smaller sections and moving cyclists from the road to more dedicated infrastructure first.

The recommendations that follow from this study are based on the found relations of which most are completely new. It would be useful to also evaluate whether the same relations are present on other road types and other cities or countries and find out whether they are universal or only local, for example due to different mobility environments or cultural differences. Further research will also help decide whether a bicycle facility of a certain width could better have its obstacles from the side in a certain zone removed and its width kept the same, or its width increased and the obstacles untouched. From the research questions asked for this study, that finding was not an urgent matter, however.

Concluding, this study contributed more empirical evidence to road safety research, that predominantly focused on matters such as the heavier users of infrastructure: motor vehicles. Especially in other countries, cyclist research and developments lag behind, so it is worthwhile for the Netherlands to take the lead. Relations on street level also help focus on anomalies or peculiarities on detailed sections of streets, that could be hot spots for bicycle crashes. The field of cyclist road safety or the more general active mode road safety is not fully explored yet, and more is certainly still to be done.

References

AASHTO. (2018). A Policy on Geometric Design of Highways and Streets: The Green Book. *American Association of State Highway and Transportation Officials, Washington, DC*, *7*, 1048.

Annema, J., & van Wee, B. (2009). *Verkeer en vervoer in hoofdlijnen (Basics of traffic and transport)*. Bussum: Uitgeverij Coutinho.

ANWB. (n.d.). *Sterrensysteem voor veiliger fietsverkeer (Star system for safer bicycle traffic).* Retrieved 2019-07-12, from `https://www.anwb.nl/belangenbehartiging/verkeer/verkeersveiligheid/sterrensysteem-voor-veiliger-fietsverkeer`

ANWB. (2020). *Dagelijkse drukke trajecten ochtend en avondspits (Daily busy morning and evening peak routes).* (Accessed June 1, 2020, from https://www.anwb.nl/verkeer/nederland/verkeersinformatie/dagelijkse-drukke-trajecten)

Bagui, S., & Mehra, K. (2019). On the Convergence of Negative Binomial Distribution. *American Journal of Mathematics and Statistics*, *9*, 44–50. doi: 10.5923/j.ajms.20190901.06

Bax, C. A., van Petegem, J. W. H., & Giesen, M. (2014). *Passen gemeenten de Ontwerpwijzer Fietsverkeer toe?: gebruik van de richtlijnen voor fietsinfrastructuur en factoren die dit beïnvloeden (Do municipalities apply the Design Guide for Bicycle Traffic?: use of the guidelines for cycling infrastructure and influencing factors).* Den Haag, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.

Bijleveld, C. C. J. H., & Commandeur, J. J. F. (2012). *Multivariate analyse: een inleiding voor criminologen en andere sociale wetenschappers (Multivariate analysis: an introduction for criminologists and other social scientists).* Den Haag, Boom Lemma Uitgevers, 2012, XXII+ 417 p., ref.; 3e gew. druk.

Buehler, R., & Pucher, J. (2012). Walking and cycling in Western Europe and the United States: trends, policies, and lessons. *Tr News*(280).

Burden, D. (2006). Urban street trees: 22 benefits specific applications. *Glatting Jackson and Walkable Communities Inc.*

Cambridge University Press. (2020). *Cambridge Dictionary.* Retrieved from `https://dictionary.cambridge.org/` (Entries for "Curb", "Obstacle", "Infrastructure" and "Statistics")

Castro, A., Kahlmeier, S., & Gotsch, T. (2018). Exposure-adjusted road fatality rates for cycling and walking in European countries..

Copenhagenize Design Company. (2020). *2019 Copenhagenize Index.* Retrieved from `https://copenhagenizeindex.eu/`

CROW. (2017). *Ruimte, mobiliteit, stedenbouw en verkeer: Licht en ruimte - openbare verlichting in de omgeving (Space, mobility, urbanism and traffic: Light and space - public lighting in the environment).* Retrieved 2020-03-02, from `https://kennisbank.crow.nl/Kennismodule`

CROW. (2018a). *Fietsverkeer: Ontwerpwijzer bruggen voor langzaam verkeer (Bicycle traffic: Design guide for bridges for slow traffic).* Retrieved 2020-02-05, from `https://kennisbank.crow.nl/Kennismodule`

CROW. (2018b). *Fietsverkeer: Ontwerpwijzer fietsverkeer (Bicycle traffic: Design guide for bicycle traffic).* Retrieved 2020-02-05, from `https://kennisbank.crow.nl/Kennismodule`

CROW. (2018c). *Wegontwerp bubeko met HWO: Handboek wegontwerp 2013 - Basiscriteria (Road design outside built-up areas: Manual for road design 2013 - Basic criteria).* Retrieved 2020-03-02, from `https://kennisbank.crow.nl/Kennismodule`

Dawson, J. (2014, 03). Moderation in management research: What, why, when, and how. *Journal of Business and Psychology*, *29*. doi: 10.1007/s10869-013-9308-7

DiGioia, J., Watkins, K. E., Xu, Y., Rodgers, M., & Guensler, R. (2017). Safety impacts of bicycle infrastructure: A critical review. *Journal of safety research*, *61*, 105–119.

Dziak, J. J., & Henry, K. L. (2017). Two-part predictors in regression models. *Multivariate behavioral research*, *52*(5), 551–561.

Fabriek, E., De Waard, D., & Schepers, J. P. (2012). Improving the visibility of bicycle infrastructure. *International journal of human factors and ergonomics*, *1*(1), 98–115.

Fietsersbond, SWOV and VVN. (2020). *Verkeersveiligheidsvergelijker (Road Safety Comparison Tool).* Retrieved from `https://verkeersveiligheidsvergelijker.nl/`

Gemeente Amsterdam. (2016, September). *Meerjarenplan Verkeersveiligheid 2016-2021 (Multi-year Road Safety Plan 2016-2021).*

Handy, S., Van Wee, B., & Kroesen, M. (2014). Promoting cycling for transport: research needs and challenges. *Transport reviews*, *34*(1), 4–24.

He, H., Wang, W., Chen, D.-G., & Tang, W. (2015). On the Effects of Structural Zeros in Regression Models. In *Innovative statistical methods for public health data* (pp. 97–115). Springer.

Higgins, J. P., Li, T., & Deeks, J. J. (2019). Choosing effect measures and computing estimates of effect. *Cochrane Handbook for Systematic Reviews of Interventions*, 143–176.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences (vol. 663).* Houghton Mifflin College Division.

Hodges, L. (1994). Common Univariate Distributions. *Methods of Experimental Physics*, *28*, 35–61.

Holsclaw, T., Hallgren, K. A., Steyvers, M., Smyth, P., & Atkins, D. C. (2015). Measurement error and outcome distributions: Methodological issues in regression analyses of behavioral coding data. *Psychology of addictive behaviors*, *29*(4), 1031.

Hutchinson, M. K., & Holtman, M. C. (2005). Analysis of count data using Poisson regression. *Research in Nursing & Health*, *28*(5), 408–418.

Iacobucci, D., Schneider, M. J., Popovich, D. L., & Bakamitsos, G. A. (2016). Mean centering helps alleviate “micro” but not “macro” multicollinearity. *Behavior research methods*, *48*(4), 1308–1317.

Iacobucci, D., Schneider, M. J., Popovich, D. L., & Bakamitsos, G. A. (2017). Mean centering, multicollinearity, and moderators in multiple regression: The reconciliation redux. *Behavior research methods*, *49*(1), 403–404.

iRAP. (2010). *Pedestrian Fencing.* Retrieved from `http://toolkit.irap.org/default.asp?page=treatment&id=56`

Kennisnetwerk SPV. (2018). *Veilig van deur tot deur: het Strategisch Plan Verkeersveiligheid 2030 (Safely from door to door: the Strategic Plan for Road Safety 2030).*

Kennisnetwerk SPV. (2019). *Investeren in verkeersveiligheid: vijf maatregelen om het fundament op orde te krijgen (Investing in traffic safety: five measures for a right foundation).*

Kerssies, J. W. (n.d.). *Veel verkeersborden zorgen voor verwarring en onveilige situaties (Many traffic signs cause confusion and unsafe situations).* Retrieved from `https://www.verkeerskunde.nl/artikel/veel-verkeersborden-zorgen-voor-verwarring-en-onveilige-situaties`

Klok Real Estate. (n.d.). *Stadsdelen Amsterdam (City districts of Amsterdam).* Retrieved 2020-02-19, from `https://www.klokrealestate.nl/124-8bcc/klok-real-estate-informatie-stadsdelen-amsterdam`

KNMI. (2018). *Burgerlijke schemering (Civil twilight).* Retrieved from `https://www.knmi.nl/kennis-en-datacentrum/achtergrond/klimatologische-brochures-en-boeken`

KNMI. (2019). *Tijden van zonsopkomst en -ondergang 2020 (Times of sunrise and sunset 2020).* Retrieved from `https://www.knmi.nl/kennis-en-datacentrum/achtergrond/klimatologische-brochures-en-boeken`

Kocur-Bera, K., & Dudzinska, M. (2015). Roadside vegetation — the impact on safety. *Proc. Intl. Sci. Con.(13)*, 594–600.

Leibowitz, H. W. (1987). Ambient illuminance during twilight and from the moon. In *Night vision: Current research and future directions. symposium proceedings* (pp. 19–22).

McCullagh, P., & Nelder, J. (1989). *Generalized linear models second edition chapman & hall.* London.

Nieuws, B. O. (2019). *Nieuwe fietsroutes moeten Amsterdam ontlasten.* Retrieved from `https://www.blikopnieuws.nl/nieuws/272595/nieuwe-fietsroutes-moeten-amsterdam-ontlasten.html` (Cover image)

NS. (2020). *Wanneer reis je met korting? (When do you travel at a discount?).* (Accessed June 1, 2020, from https://www.ns.nl/uitgelicht/wanneer-reizen-met-voordeel/wanneer-reist-u-met-korting.html)

Ormel, W., Klein Wolt, K., & Den Hertog, P. (2008). Enkelvoudige fietsongevallen (Single-bicycle crashes). *Amsterdam: Stichting Consument en Veiligheid.*

Othman, S., Thomson, R., & Lannér, G. (2009). Identifying critical road geometry parameters affecting crash rate and crash type. In *Annals of advances in automotive medicine/annual scientific conference* (Vol. 53, p. 155).

O’Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & quantity*, *41*(5), 673–690.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *50*(302), 157–175.

Raihan, M. A., Alluri, P., Wu, W., & Gan, A. (2019). Estimation of bicycle crash modification factors (CMFs) on urban facilities using zero inflated negative binomial models. *Accident Analysis & Prevention*, *123*, 303–313.

Reynolds, C. C. O., Harris, M. A., Teschke, K., Cripton, P. A., & Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. *Environmental health*, *8*(1), 47.

Rijkswaterstaat. (n.d.). *Verkeersveiligheid en ongevallencijfers (Road safety and crash figures).* Retrieved 2020-02-26, from `https://www.rijkswaterstaat.nl/wegen/wegbeheer/onderzoek/verkeersveiligheid-en-ongevallencijfers/index.aspx`

Schaap, T. W., Harms, L. W. J., Kansen, M., & Wüst, H. (2015). *Fietsen en lopen: de smeerolie van onze mobiliteit (Cycling and walking: the lubricant of our mobility).* Kennisinstituut voor Mobiliteitsbeleid.

Schepers, P. (2008). De rol van infrastructuur bij enkelvoudige fietsongevallen (The role of infrastructure in single-bicycle crashes). *Directoraat-Generaal Rijkswaterstaat, Dienst Verkeer en Scheepvaart DVS, Delft*.

Schepers, P., & den Brinker, B. (2011). What do cyclists need to see to avoid single-bicycle crashes? *Ergonomics*, *54*(4), 315–327.

Schepers, P., Hagenzieker, M., Methorst, R., Van Wee, B., & Wegman, F. (2014). A conceptual framework for road safety and mobility applied to cycling safety. *Accident Analysis & Prevention*, *62*, 331–340.

Schepers, P., & Klein Wolt, K. (2012). Single-bicycle crash types and characteristics. *Cycling Research International*, *2*(1), 119–135.

Schepers, P., Stipdonk, H., Methorst, R., & Olivier, J. (2017). Bicycle fatalities: trends in crashes with and without motor vehicles in The Netherlands. *Transportation research part F: traffic psychology and behaviour*, *46*, 491–499.

Stipdonk, H. L., & Bos, N. (2014). *Verkeersveiligheidsanalyse gemeente Amsterdam: relevante doelgroepen, ontwikkelingen en aandachtspunten (Road safety analysis for the municipality of Amsterdam: relevant target groups, developments and points of attention).* Den Haag, SWOV.

SWOV. (n.d.). *Organisatie (Organisation).* Retrieved 2019-12-05, from `https://www.swov.nl/over-swov/organisatie`

SWOV. (2012). *Risico in het verkeer (Risks in traffic).* (SWOV factsheet)

SWOV. (2016). *Data sources: Comprehensive overview.*

SWOV. (2017). *Principes voor veilig wegontwerp (Principles for safe road design).* (SWOV factsheet)

SWOV. (2018). *Duurzaam Veilig Wegverkeer (Sustainable Road Safety).* Retrieved 2019-12-09, from `https://duurzaamveiligwegverkeer.nl/`

SWOV. (2019). *Publicaties (Publications).* Retrieved 2019-12-09, from `https://www.swov.nl/publicaties/swov/type/publication/updated/2019/publication{_}type/rapport`

VeiligheidNL. (2019). *Aantal fietsers met ernstig letsel op de Spoedeisende Hulp gestegen (Number of cyclists with serious injuries to the emergency room increased).* Retrieved 2019-12-05, from `https://www.veiligheid.nl/verkeer/actueel/aantal-fietsers-met-ernstig-letsel-op-de-spoedeisende-hulp-gestegen`

Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods*, *17*(2), 228.

Wanvik, P. O. (2009). Effects of road lighting: an analysis based on Dutch accident statistics 1987–2006. *Accident Analysis & Prevention*, *41*(1), 123–128.

Wardlaw, M. J. (2014). History, risk, infrastructure: perspectives on bicycling in the Netherlands and the UK. *Journal of Transport & Health*, *1*(4), 243–250.

Weijermars, W. (2019). *Monitor Verkeersveiligheid 2019: effectieve maatregelen nodig om het tij te keren (Road Safety Monitor 2019: effective measures are required to stem the tide).* Den Haag, SWOV.

Weijermars, W., Van Schagen, I., Aarts, L., Van Petegem, J., Van Stipdonk, H., & Wijnen, W. (2018). *Hoe verkeersveilig kan Nederland zijn in 2030?: mogelijkheden voor reductie in aantallen verkeersslachtoffers (How safe can the Netherlands be in 2030?: possibilities for traffic victim reductions).* Den Haag, SWOV.

Wijlhuizen, G. J., van Petegem, J. W. H., Goldenbeld, C., van Gent, P., de Bruin, J., de Commandeur, J. J. F., & Kars, V. (2016). *Doorontwikkeling CycleRAP-instrument voor veiligheidsbeoordeling fietsinfrastructuur: doelmatigheid handmatige intensiteitsmetingen, betrouwbaarheid beoordelingen infrastructuur en validiteit* *van het CycleRAP-instrument (Further development of the CycleRAP instrument for safety assessment of cycling infrastructure: effectiveness of manual intensity measurements, reliability of infrastructure assessments and validity of the CycleRAP instrument)).* Den Haag, SWOV.

Wijlhuizen, G. J., van Petegem, J. W. H., Schermers, G., de Bruin, J., de Commandeur, J. J. F., & Kars, V. (2017). *Ontwikkeling Netwerk Safety Index gemeente Amsterdam (Network Safety Index development for the municipality of Amsterdam).* Den Haag, SWOV.

Wood, A. G., Mountain, L. J., Connors, R. D., Maher, M. J., & Ropkins, K. (2013). Updating outdated predictive accident models. *Accident Analysis & Prevention*, *55*, 54–66.

World Health Organization. (2009). *Global status report on road safety: time for action.*

World Health Organization. (2019). *Decade of Action for Road Safety 2011-2020.* Retrieved 2019-12-11, from `https://www.who.int/roadsafety/decade_of_action/en/`

Yan, J., Guszcza, J., Flynn, M., & Wu, C.-S. P. (2009). Applications of the offset in property-casualty predictive modeling. In *Casualty actuarial society e-forum, winter 2009* (p. 366).

A

Paper

Relations between the obstacle space of cycling infrastructure and bicycle crashes: An analysis of Amsterdam

Guus van Weelderen

Delft University of Technology
Faculties of Civil Engineering and Geosciences & Technology, Policy and Management

Abstract

The objective of this paper is to strengthen literature about cyclist safety and help prevent bicycle crashes in the future. It looks into the gap about the relation between the obstacle space of cycling infrastructure (that involves the space that is available for cyclists to avoid crashes with the obstacles therein) and the visibility thereof, and bicycle crashes. Next to a brief literature research, the main method used to answer this question is modelling of generalised linear models with a negative binomial distribution, that link independent variables concerning the obstacle space to the dependent count variable crashes. Results show that the relation is apparent in data concerning 50 km/h streets in the municipality of Amsterdam: infrastructure on a street level can indeed predict bicycle crashes on the same level. Fewer obstacles next to the road and higher widths were found to lead to fewer crashes. For lanes and one-way tracks holds that an increase in width of smaller sections has a bigger impact. The effect of obstacles in the shoulder also reduces at higher widths. A significant relation for light conditions was not found. Further research should look into different types and clustering of obstacles and improve the light analysis. Recommendations include providing dedicated cycling infrastructure with sufficient widths and few obstacles, and better registration of cyclist crashes. The study demonstrates that infrastructure is an important factor for the occurrence of bicycle crashes, that should certainly not be disregarded alongside the road user and vehicle. It hopes to push infrastructure policies for cycling more on the agenda.

Keywords: active modes, bicycle, accidents, swerving, obstacle-free zone, shoulder, transition, light conditions, generalised linear regression, negative binomial distribution

1. Introduction

1.1. Research context

Cycling is beneficial to both health and the environment, and is relatively cheap (Handy, Van Wee, & Kroesen, 2014). It also is an important mode of transport in the Netherlands. About a quarter of the trips in the country are made by bicycle (Buehler & Pucher, 2012). The average exposure in yearly travelled kilometres per person for cycling is among the highest in the world, and per hundred million kilometres cycled, 0.8 fatalities are registered. From the five-year averages of cyclists deaths in traffic for several European countries in Figure 1 follows that a higher exposure for cycling per person per year in a country suggests a lower risk of getting killed in a crash. The Netherlands places itself at the right-hand side of the graph, and thus already performs relatively well in this aspect (Castro, Kahlmeier, & Gotsch, 2018). Still, the risk per kilometre travelled to die in traffic in 2009 was about a factor 5 higher for cyclists than for someone in a car, and to get severely injured about 40 (SWOV, 2012).

Around a third (228 people) of traffic deaths and over 60% of the 123,000 traffic victims on the emergency department in 2018 were cyclists. The number of cyclists that was treated for severe injuries even increased to 46,800, which is nearly a third higher than in 2009 (VeiligheidNL, 2019). From 1996 to 2014, the number of cyclist deaths without interference from motor vehicles increased by 7% per year, while those with motor vehicles decreased by 3.8% per year. Despite that, the total number of cyclist deaths remained stable during that period, while the annual road deaths through other modes more than halved (Schepers, Stipdonk, Methorst, & Olivier, 2017). The total trend appears to have stopped in recent years, going up again from 2013 (see Figure 2). Based on a continuation of the current transport policies, a forecast for the Netherlands assumes around the same shares of deaths and injured for cyclists in 2030 (Weijermars et al., 2018). Since it is not feasible for the upward or even stagnating trend of traffic deaths to continue, with high cyclist shares, things need to change. A reduction in the number of traffic deaths is already part of several plans, such as the Meerjarenplan Verkeersveiligheid 2016-2021 ("Multi-year Road Safety Plan 2016-2021") by the municipality and the Strategisch Plan Verkeersveiligheid 2030 ("Road Safety Strategic Plan 2030") by the Dutch government and other organisations. Like a plan by the European Commission, the latter strives to achieve zero traffic deaths in 2050 (Gemeente Amsterdam, 2016; Kennisnetwerk SPV, 2018). It therefore is worthwhile to look for different ways to make cycling safer and reduce the involved risks.

1.2. Research problem

One way to make cycling safer is by focusing on the infrastructure. For this, general relations between infrastructure

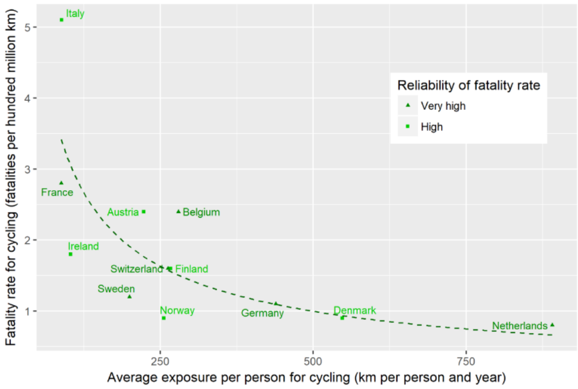


Figure 1: Fatality rate versus exposure (Castro et al., 2018)

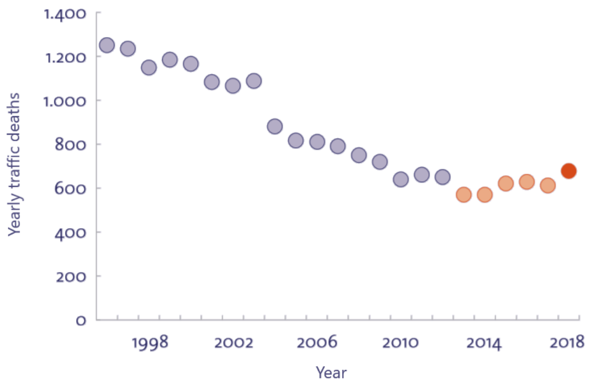


Figure 2: Yearly traffic deaths (Weijermars, 2019)

and bicycle crashes need to be known: which elements improve safety, and which decrease it? This study focuses on the obstacle space for cyclists, which was defined here as the space that is available for cyclists to avoid crashes with the obstacles therein, and the visibility of this space. Infrastructure and its relation to crashes have been studied numerous times in the past, but the aforementioned does not seem to be part of that.

Broadly speaking, facilities for cyclists (Reynolds, Harris, Teschke, Cripton, & Winters, 2009), and bicycle streets (priority for cyclists over motorised vehicles) and bicycle tracks (separated from roads) have been found to reduce crashes and injuries for cyclists. Moreover, cycling at higher volumes appeared to positively influence safety. This is known as the "safety in numbers" phenomenon (DiGioia, Watkins, Xu, Rodgers, & Guensler, 2017). Next to sufficient street lighting, paved surfaces and lower-sloped roads (Reynolds et al., 2009), improvements to the general road safety are beneficial as well (Wardlaw, 2014). The quality of infrastructure elements, or design in general, also is an important point: not merely the presence of elements is of interest, but also how they are implemented. Differences between countries furthermore make that the aforementioned risk-reducing or increasing factors may not hold everywhere (Schepers, Hagenzieker, Methorst, Van Wee, & Wegman, 2014).

In the Netherlands, the risk of single-bicycle crashes due to infrastructure did not differ significantly between road types (bicycle tracks, bicycle lanes, etc.). About half of these crashes were caused by infrastructure-related factors. The influence of the width of bicycle lanes was mentioned as a future research possibility (Schepers, 2008). Infrastructure-related single-bicycle crashes in a later study were caused by collisions with obstacles, riding off the road, skidding and stability issues. Other causes were related to the cyclist, their bicycle or external factors. Specifically, in situations where cyclists had right of way, on intersections with two-way bicycle tracks that were marked and coloured red, crashes happened more. On the other hand, raised bicycle crossings and other motor vehicle speed-reducing measures reduced the risk (Schepers & den Brinker, 2011; Schepers & Klein Wolt, 2012). A later study by Wijlhuizen et al. (2016) found no relation between obstacles (in a compound score) and the quality of cycling infrastructure, including width, with bicycle crashes. They state it should be investigated in other locations. Higher cycling and motor vehicle intensities, and higher intersection densities both related to a higher number of bicycle crashes. The severity of injuries did not make a difference, so it was recommended to only use one model with the general crash numbers.

Road lighting decreased crashes with injuries by 50% during darkness and during twilight by 2/3 of that. In rural areas, lighting was more effective. It also had more effect on pedestrians, cyclists and moped users compared to other motor vehicles. Infrastructure characteristics were not taken into account (Wanvik, 2009). Single-bicycle crashes were more frequent under conditions with less light, but insignificant, possibly due to older cyclists cycling less in darkness. Bad visual characteristics lead to collisions with bollards or narrowings or riding off the road, recommended to be prevented by edge markings and better visible bollards. Looking into whether motorists would use cycle tracks if bollards would be removed, and limiting consequences of bicycle crashes by designing bicycle facilities wide enough, with a small height difference with the sides is proposed (Schepers & den Brinker, 2011). Suggestions for improvement of visibility were previously given in Fabriek, De Waard, and Schepers (2012) as well.

With cyclists and single-bicycle crashes taking up the highest shares in traffic deaths and severe injuries, and infrastructure being the most common cause of single-bicycle crashes, its importance is clear. However, a research gap indeed still exists. None of the studies specifically looked into the crash risk of obstacles, width and light conditions for cyclists and how this risk changes in different combinations. When cyclists are riding their bicycle and face an obstacle, the width available namely limits their manoeuvring space. If the road section they ride on is several metres wide, there is a higher chance they can swerve around it or otherwise avoid a collision without loss of balance than when it is only half a metre wide. During the day, it is also easier to see obstacles. In conjunction, say obstacles on a route with narrower road sections during the night, it can be much harder to avoid a crash. This gap will thus be addressed in this research. While it is assumed that light has an impact on crashes, this topic will only be briefly touched in this study due to scoping.

1.3. Research objective and questions

To test if any of the previous relations between the obstacle space and bicycle crashes are present, a real-world application is empirically tested. The Dutch national institute for road safety research SWOV has conducted studies into infrastructure and road crashes involving both all modalities and cyclists in the recent past, partly commissioned by the municipality of Amsterdam (Wijlhuizen et al., 2016, 2017). As such, data are available. With the city being the capital and biggest city in the country that receives a lot of tourists, a lot of cyclists make use of the infrastructure. Of the four biggest cities in the Netherlands (Amsterdam, Rotterdam, The Hague and Utrecht), Amsterdam comes second when looking at the trip modal split, with 33% being made by bicycles (Schaap, Harms, Kansen, & Wüst, 2015). In the latest Copenhagenize Index from 2019, that ranks cities in the world according to their bicycle-friendliness, Amsterdam places second as well, below Copenhagen (Copenhagenize Design Company, 2020). The city of Amsterdam is thus deemed a good case study again.

The objective of this study is to address the gap in literature of the relation between obstacles, width, light conditions and their interactions with bicycle crashes in Amsterdam, as to help researchers, municipalities and other organisations understand more about cyclists and cycling and help the Netherlands become safer and more bicycle-friendly. For example, any relations found can be put into practice by adjusting unsafe infrastructure on bicycle routes and used as further evidence for projects like CycleRAP, a scoring method for cycling infrastructure based on the probability of a crash (Wijlhuizen et al., 2016). The main question that was formulated is the following:

To what extent can the obstacle space of cycling infrastructure and visibility thereof explain bicycle crashes in Amsterdam?

Sub-questions to support this main question are related to the general characteristics of the obstacle space of cycling infrastructure, the relation between obstacles on and around cycling infrastructure, cycling infrastructure width, light conditions, and combinations of the aforementioned and bicycle crashes, and measures that can be taken from these relations to improve cycling infrastructure in the Netherlands and reduce crash risk. The coming sections of this paper will answer these questions. Section 2 will discuss the question about obstacles and width, and end with a conceptual model. In Section 3, information about data gathering and processing and the general methodology for the analysis will be explained, consecutively elaborating on the data and which variables are created. The next set of questions are answered in Section 4 by taking a look at the modelling results. The study concludes with shortcomings and further steps.

2. Literature research

2.1. Element characteristics

To interpret the results of the coming analyses, it is useful to take a look at the underlying elements related to obstacle avoidance and how these can lead to crashes. In the following paragraphs, a look will be taken at the types of obstacles and ranges of widths, and determining factors for implementation and crash risk. In several cases, guidelines by the technology platform for transport, infrastructure and public space in the Netherlands are cited. The guidelines for cycling were first drafted in the nineties, based on experience from work groups based on consensus, logical reasoning and simply adding measurements, and evidence from previous studies. Bicycle policies are often local, which means the local situation needs to be taken into account. Therefore, the guidelines are supplemented with arguments, experiences, ideas and tips (CROW, 2018b). The guidelines also do not have a legal status, so it is allowed to deviate from them. In a study among several municipalities, 56% of the respondents mention they barely use the Design Manual for Bicycle Traffic. Lack of space, costs and awareness of the guidelines are reasons why the recommendations for the width of bicycle tracks, obstacle-free zone and profiled road markings at bollards are not followed (Bax, van Petegem, & Giesen, 2014).

2.1.1. Obstacles

Typical obstacles considered in a survey by Ormel, Klein Wolt, and Den Hertog (2008) and used by Schepers and Klein Wolt (2012) are lamp posts, traffic signs, bollards, fences, walls, kerbs and trees, along with the options "Other" and animals that are not infrastructure-related. Under "Other", a type of obstacles to consider are parked cars. From that study, distinctions made here are whether it is located on or off the road (the "site"), whether the obstacle is fixed or dynamic, i.e. is it always there or only occasionally (the "state"), and whether it is a point or line obstacle, i.e. is it concentrated on a single spot or a longer stretch (the "shape")? These types already cover most obstacles. Crash risk for those is related to collision, distraction, blocking vision and collision as a result of the latter two. An overview of all functions and the classification is shown in Table 1.

2.1.2. Width

The width of a road for cyclists depends on several factors. Firstly, it is important to take the dimensions of the users and any vehicles they use into account. As cyclists and bicycles come in many forms, it is not useful to design infrastructure for the average bicycle, and the legal requirements serve as a better basis. These state that a bicycle can be 0.75 metres wide at maximum, and mopeds 1.10 metres. Next to this, some extra space is added. Since a bicycle just has two wheels, the rider namely needs to keep their balance and as a result makes small corrections while cycling. This results in a sway, that makes that cyclists do not ride in a perfectly straight line. These come together in the free-space profile, that is shown in Figure 3 and is the space around a cyclist a designer needs to bear in mind (CROW, 2018b).

Another factor for the width of paving is the fear of obstacles on the sides, including other cyclists, with a safety distance of at least 0.25 metres. With every cyclist in a cross-section that gets added, the width ideally thus needs to be increased

Table 1: Overview obstacles with function and classification

Obstacle	Function	Classification		
		Site	*State*	*Shape*
Lamp post	Provide light	Off	Fixed	Point
Traffic sign	Provide information	Off	Fixed	Point
Bollard	Control access for motor vehicles	On/off	Fixed	Point/line
Fence	Define barrier, control access and guard	Off	Fixed	Line
Wall	Define barrier, support structures and guard	Off	Fixed	Line
Kerb	Define barrier, control access and more	On/off	Fixed	Line
Tree/vegetation	Define barrier, reduce speeds and more	Off	Fixed	Point/line
Parking place	Store vehicle	On/off	Dynamic	Line

by at minimum 1 metre. The intended use (either one or two directions, and possibilities for overtaking) and the expected traffic intensities are of interest too. For a bicycle-moped track, 0.5 metres extra is usually added (CROW, 2018a). Depending on a combination of the aforementioned and the type of bicycle facility (bicycle tracks, bicycle-moped tracks, (non-designated) bicycle lanes and bicycle streets), the minimum width values per direction for cyclists typically range from 2 to 5 metres, with a special mention for 1.5 metres for solitary bicycle tracks at low intensities and with easily passable kerbs or shoulders and at least 1 metre on the carriageway.

2.2. Conceptual model

To prepare for the methodology, the previous points should be part of a greater scheme that shows the hypothesised relations. For this, a conceptual model (Figure 4) is adapted and adjusted from one in Schepers et al. (2014), that combined the passenger transport model by Annema and van Wee (2009) with the traffic safety pillars by Othman, Thomson, and Lannér (2009). Along with external factors, these pillars are also the categories of single-bicycle crash causes found by Schepers and Klein Wolt (2012). The influences and interactions of the different parts on each other are illustrated by arrows. The boxes in blue stand for the focus of this study. Rules and regulations, which were not included in the analyses, are imperatively considered at the end. The different road users each have their own capabilities and make decisions that influence the risk. Characteristics such as age and sex are assumed to be of influence on these risks. Since the scope of this project limits itself around infrastructure, only the presence of users on the roads is used. Next to the users, the vehicles they drive can influence the risk. These can differ per type and number of vehicles that drive on a road, but also the state they are in. While it would be valuable to also test for data surrounding vehicle, these are not present and will thus not be taken into account, apart from the fact that a bicycle needs to have been involved for a crash to be called a bicycle crash.

The infrastructure they ride on comes in different forms. For this study, obstacles, width and light conditions are of main importance. Any other factors that are not related to those factors later on are in the category "Other". Internally, these elements might affect each other, such as the width and obstacles. Both direct and indirect effects of the infrastructure characteristics are expected. More obstacles on and next to the road are assumed to increase the risk of colliding with one, and the wider the bicycle road infrastructure, the more space is available for cyclists to avoid crashes, for example due to collisions with obstacles. With more natural light and street lighting available, people feel safer and are less afraid to ride faster. As such, more crashes may happen. Indirect relations with bicycle crashes are assumed to be between these obstacles, width and light conditions. If a street has a

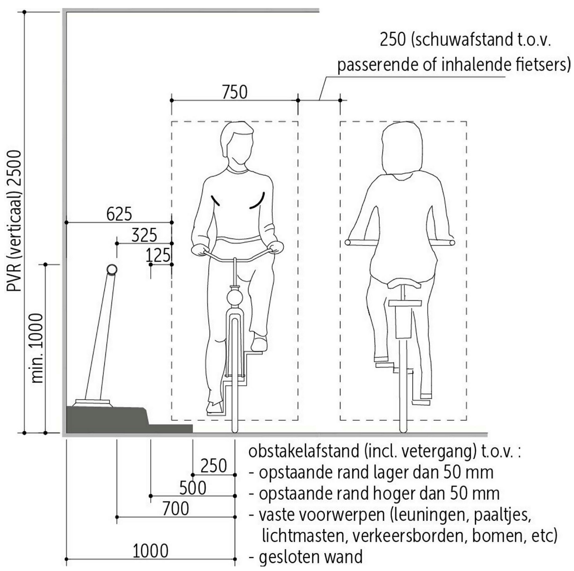


Figure 3: Free-space profile (CROW, 2018b)

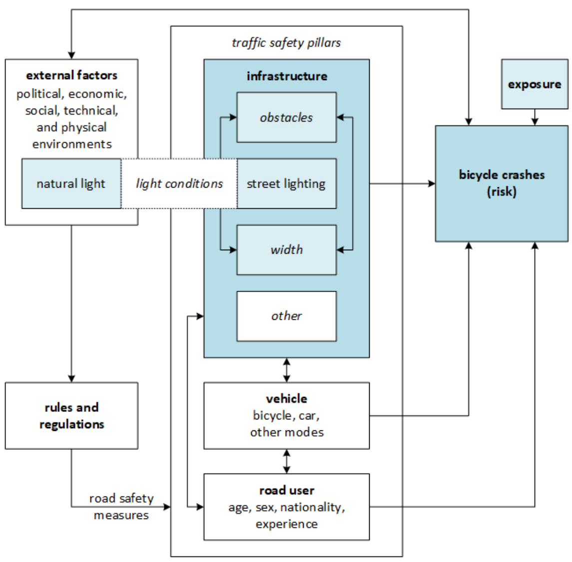


Figure 4: Conceptual model

certain number of obstacles but has better light conditions or gets widened, these obstacles might be less of a problem. A certain width might also not be as effective in reducing crash risk once the light conditions are worse. These direct and indirect relations will be tested.

There are also interactions between pillars. Between the road user and a vehicle, an elderly person on a regular bicycle or an e-bike may for example not yield the same risk. Someone can also have plenty of experience with a certain vehicle or fail to steer away when needed, or certain vehicles might be more forgiving in the case of a crash. Between the road user and infrastructure, you can think of someone with either good or bad sight: there is a chance the latter overlooks a bollard (related to the research by Schepers and den Brinker (2011)). Between infrastructure and a vehicle, some vehicles are not allowed to ride on a road stretch or better able to handle less smooth surfaces.

Next to the pillars, there are external factors that can impact them, and thus the risk of bicycle crashes. Among these factors are the political, economic, social, technical and physical environments, that can lead to rules and regulations (road safety measures) or change aspects of the pillars and/or risks. If a crash happens, this can feed back to the external factors. Things such as climate change can also pressure governments to adapt or revise regulations. People can become more risk-averse when travelling. Surroundings, such as a school or shopping area, and the weather or time of year are examples of the physical environment that might affect risks. In particular, the total duration of daylight will be looked into here. It makes that obstacles are less visible. Lastly, exposure needs to be taken into account: the more someone exposes themselves to a street and other people (as expressed in the street length and intensities), the higher the risk they will get involved in a crash.

3. Methodology

3.1. Data collection and assessment

To fully assess the effect of the obstacle space and study the relation between infrastructure and the risk they pose on bicycle crashes, it is important to gather data about the crashes, exposure (such as volumes) and roadway characteristics. For the first, mostly underreporting is a problem (DiGioia et al., 2017). Since SWOV already conducted several studies in which on a large scale data on Amsterdam are obtained and the topic was proposed by them, most data were available in the data set provided at the start of the study. The study focuses on crashes on distributor roads with a speed limit of 50 km/h in Amsterdam. In 2015, the infrastructure of the Centrum, Oost, Zuid, West and Nieuw-West districts were assessed. In 2016, the remaining and more outer districts of Noord, Zuidoost, Westpoort followed (Wijlhuizen et al., 2017). Figure 5 shows the location of these areas. The streets are also depicted in Figure 6. The blue roads were part of the first group, and the red ones of the second.

Some filtering of the data was required, since not all were deemed usable. After several steps, from the infrastructure data set with 710 streets, only 337 streets remained. Any streets that in the end could not have any crash numbers extracted, including zero, were removed. Streets that did not have a speed limit of 50 km/h for 90% or more of their length were excluded as well. This was taken from Wijlhuizen et al. (2017). Most severe crashes namely happen on these streets in cities (Stipdonk & Bos, 2014). By filtering out streets with a length shorter than 150 metres (an arbitrary number taken from the aforementioned study), several streets with special layouts, like bridges and squares, were eliminated too. Lastly, any streets still present that were missing data for the variables mentioned later were removed. Incorrect or irrelevant data could still be present due to human error. Due to the amount of work to be done, most entries were deemed fine, since it concerns at most a dozen incorrect values per category and not all variables will be used. Any elements that could not be determined at the time were not revisited and corrected.

3.1.1. Crash data

Sources for the crash data are ambulance data. Data from the National Institute for Public Health and the Environment (RIVM) from 2009 to 2012 were used in Wijlhuizen et al. (2017). To this, data from the municipality of Amsterdam about crashes from 2014 to 2016 are added. Since not all

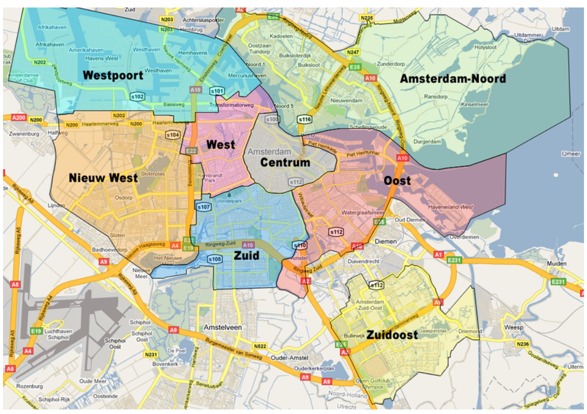


Figure 5: Districts of Amsterdam (Klok Real Estate, n.d.)

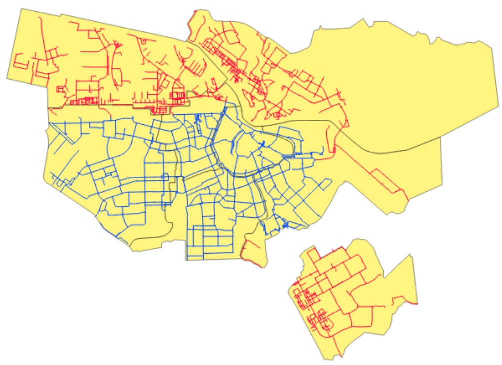


Figure 6: 50 km/h roads in Amsterdam

of these data are logged on exact 4- and 6-digit postal code levels, it is not possible to exactly link certain environments to crashes and the analysis will be done on street level. In general, of the filtered set, just over 20% of the streets had no crashes reported. Crashes where the mode was registered account for about 42% (5989 crashes). Of these, almost 45% of the crashes involved the bicycle (2707 crashes).

3.1.2. Exposure data

Exposure data gives a general idea about how long or with how many people someone is exposed to the variables to be studied, and that may result in a crash. In this case, these are the street length and intensities. The longest street is just above 5300 metres, and the 30 streets between 150 and 200 metres account for about 9% of the total streets. This share generally decreases as length goes on, with the mean being around 937 metres. Furthermore, the available intensities encompass both motor vehicle intensities and cyclist intensities. The first are taken from the traffic model for 2010 by the municipality of Amsterdam in annual average daily traffic. The bicycle intensities are averages of the intensities gathered during the Fietstelweken ("Bicycle Count Weeks") in September in 2015 and 2016 (Wijlhuizen et al., 2017). According to the traffic model, on average most streets in the filtered set (just over one quarter) have to cope with less than 1500 vehicles per day. The minimum is an arbitrary value of 30: it is namely not expected that a street is not used by any cars at all, and the values below 30 were generated by a model anyway. The maximum is 31906 and the mean is an average of 5797 motor vehicles per day. For cyclist intensities, it seems that during the Bicycle Count Weeks, more than forty per cent of the streets faced 30 or fewer cyclists. Again a minimum was set; this time at a value of 5. The maximum is 635 and the mean is about 87 cyclists in those weeks.

3.1.3. Infrastructure data

The roadway characteristics were extracted from images taken from CycleRAP and captured by Cyclomedia in 2015 and 2016. Only streets that the municipality of Amsterdam classified as a distributor road were considered. Even though a photo was taken every 5 metres, the characteristics were only assessed every 25 metres. Further details about the process are present in Wijlhuizen et al. (2017). The characteristics were then scaled up to street level as shares or means. As follows from the following paragraphs, the assessed streets in Amsterdam include a broad range of elements.

More than half of the transitions from the road surface to the shoulder are formed by kerbs, split about equally into impassable and passable kerbs. The next third is split between parking on and along the carriageway, and no real transition at all, namely flat. The quality of these transitions depends mostly on the height difference (also depending on cracks, holes and bumps), and whether there is a chance to lose balance. For bottlenecks, contact should be avoided, while it is uncomfortable for points of attention. 48% was found to be a bottleneck, 80% including points of attention. The most common types of shoulders are parking places, road surface such as pavement, and grass. These take up a quarter and two fifths. While not every type might seem a threat at first, still a significant share of 81% of these shoulders is labelled as a bottleneck, increasing a little to 89% including points of attention. Since points of attention of transitions and shoulders are more related to comfort and not direct danger, only bottlenecks will be considered for later. This also makes it easier to discuss implications.

Almost 6 out of 10 sections feature obstacles that are 50 centimetres or closer, even adjacent, to the road surface. Furthermore, a meagre 0.6% of the street sections contain a median (but 42 streets, or 12.5%), and nearly 0.1% contain a bollard (but 17 streets, or 5%). Given their low occurrence, this could already be a sign that it might be hard to find a relation between these obstacles and bicycle crashes from this data set. When looking at the facility types, two clearly dominate: the biggest shares are taken by the bicycle tracks (48%) and carriageway (25%). Also, around 9% of the street sections do not have lamp posts.

3.2. Variable overview

From the data sets, the variables in Table 2 and 3 were created. Scale variables concern interval and ratio levels of measurement. Nominal variables concern dummies. The obstacles mentioned before (lamp posts, traffic signs, bollards, fences, walls, kerbs, trees and parked cars) were present in some way. The categories "Type/Quality of shoulder" (all types except kerb), "Type/Quality of transition" (kerb and more), "Obstacle-free zone next to bicycle facility" (all types), "Bollard on road" (bollard) and "Median on road" (kerb) contain information about these obstacles. Since there are a lot of different types of obstacles in the shoulder and transition, of which some are more common than others, the choice was made to only focus on the overarching quality of these elements for now. Apart from some modifications to make the data suitable for the analysis model, the obstacle part as pictured in the conceptual model should be covered.

The cycling infrastructure width is also present in the data set. The annotators that extracted the width measured the smallest width at every 25-metre section and had to note the full width of two-way bicycle tracks and 1.5 metres as widths at carriageway sections. Since we want to find out the impact of the width of infrastructure on bicycle crashes that cyclists have at their disposal, the width of two-way facilities was halved and the width for cyclists on carriageways was calculated from the data as space remaining after subtracting motor vehicle widths at the available width. This could lead to negative cycling widths. In order to be able to keep enough streets in the model, and considering the similarities between or differences in several bicycle facility types, some of the types were merged or split. Non-designated bicycle lanes and bicycle lanes were combined into the simpler "lane" type, bicycle tracks were separated in the "one-way" and "two-way" track types and carriageways were kept as-is. The means of widths decrease from 2.10 m for one-way tracks, to 1.65 m for two-way tracks and lanes to 0.93 m for carriageways. Since streets typically do not contain only one type and we want to look at the impact of facility widths

Table 2: Overview of scale (interval/ratio) variables

Focus	Name	Count	Min.	Max.	Mean	Std. dev.
Objective	Cyclist crashes	337	0	105	8.03	14.18
Offset	Road length	337	150.11	5331.69	937.49	835.08
Base	Motor vehicle intensity	337	30	31906	5797.27	5639.23
	Cyclist intensity		5	635	86.94	104.31
Obstacles	Shoulder bottleneck share	337	0.00	1.00	0.83	0.21
	Transition bottleneck share			1.00	0.57	0.38
	Obstacle-free zone share - adjacent			1.00	0.25	0.32
	Obstacle-free zone share - 50 cm			1.00	0.63	0.29
	Obstacle-free zone share - 1 m			1.00	0.84	0.20
	Obstacle-free zone share - 2 m			1.00	0.94	0.12
	Median share			0.27	0.01	0.03
	Bollard share			0.20	0.00	0.01
	On-path obstacle (OPO) share			0.14	0.00	0.01
	Visibility problem share - median			0.27	0.01	0.03
	Visibility problem share - bollard			0.20	0.00	0.01
	Visibility problem share - OPO			0.14	0.00	0.01
Width	Dominating width - mean	267	-1.15	3.68	1.58	0.63
	Dominating width - std. dev.		0.00	1.90	0.15	0.24
	Narrowing share		0.00	1.00	0.02	0.07
Light	Lamp post share	337	0.00	1.00	0.92	0.16

Table 3: Overview of nominal variables

Focus	Name	Count	Frequency: 0	Frequency: 1
Obstacles	Median dummy	337	295	42
	Bollard dummy		321	16
	On-path obstacle dummy		280	57
Width	Lane dummy	267	228	39
	One-way track dummy		173	94
	Two-way track dummy		232	35

without compromising or negating width differences among streets, at the same time it was decided to look at streets with one dominating bicycle facility type; that is, more than 50% available along the street sections. This should not lead to too little streets in the model, but at the same time keep the results as pure as possible. Therefore, dummy variables for each facility type were created. Since carriageways were one of the most common types in the data set and are not specifically designed for cyclists, they are the reference group that is omitted (Dziak & Henry, 2017). The lower the dominance threshold, the more streets are included. 75% was deemed a good threshold to still have a sense of dominance, yet not reducing the data set too much.

3.3. Modelling approach

To answer the research questions, a method is needed that allows for a way to infer bicycle crash numbers from the elements in the conceptual model for which the previous variables were created, and thus streets and the environment. Statistical models are suitable for this. In statistics, regression analysis namely is used to find relations between a dependent variable and one or multiple independent factors. The use of generalised linear regression models was chosen. This type of model is more general than the classical regression models and allows other distributions than the normal distribution to be applied to the dependent variable, such as discrete distributions that are useful for crashes that are whole numbers (McCullagh & Nelder, 1989). They were previously put in a negative binomial distribution by Wijlhuizen et al. (2017). Other research has applied this negative binomial distribution for crash modelling as well (Bagui & Mehra, 2019). This distribution handles overdispersion, meaning that the variance is larger than the mean, better than what is called the Poisson distribution (Hutchinson & Holtman, 2005). After a check with the Pearson Chi-Square test for the goodness of fit (Pearson, 1900), the negative binomial distribution always performed better. The step-wise modelling approach and model formulation of Wijlhuizen et al. (2017) will be used as a starting point for the new research questions, given the similarities in research design. This might make comparing the two studies easier.

$$\begin{aligned}\ln(C) &= b_0 + \ln(L) + b_1 \ln(I_{mv}) + b_2 \ln(I_{cyc}) \\ &\quad + b_3 X_1 + b_4 X_2 + \cdots + b_z X_{z-2} + e\end{aligned} \tag{1}$$

In Equation 1, C are the bicycle crashes on a street, b_0 is the intercept, other b's are regression coefficients, L is the length of a street, I are the traffic intensities (mv for motor vehicles and cyc for cyclists) and X's represent other model variables. The length is an exposure variable, that is generally treated by either putting it in the model as another predictor variable, or by putting it in a (natural) logarithm, with coefficient 1, and adding it as another model variable called the offset. This effectively changes the crash counts to a crash density: the number of crashes per kilometre of road length on a street. The predicted outcome then also should be interpreted as the rate of crashes per road length unit (Holsclaw, Hallgren, Steyvers, Smyth, & Atkins, 2015; Yan, Guszcza, Flynn, & Wu, 2009). Since the relationship between intensities and crashes was found to be non-linear (Wood, Mountain, Connors, Maher, & Ropkins, 2013), Wijlhuizen et al. (2017) decided to put the intensities in a natural logarithm as well. This decreases the rate at which crashes occur at increasing intensities ("safety in numbers" as discussed earlier). These measures are believed to be sensible and will be maintained. After a base model without obstacle space variables, models with separate obstacle space parameters were run and model improvements were checked. This was done on a per-research question basis. The models were extended with interaction variables. As the variables appear in the world together, combined models were developed as well. Correlations were taken care of by preventing some combinations of variables from appearing in the different models.

Interpretation of regression coefficients for interaction terms is not as straightforward as for the main effects. The interaction effects are therefore usually plotted at low and high values, or simply (between) one standard deviation away from the mean if no special values are present (Dawson, 2014). Other variables, the conditional variables, should stay the same. Due to the application of mean centring, where the means of every independent variable are subtracted from the variables, means of 0 are achieved. This is to prevent multicollinearity and lower accompanying Variance Inflation Factors (VIF), with those lower than 4 preferred (Bijleveld & Commandeur, 2012; O'Brien, 2007). Conditional variables were kept at the means, as this means no extra contribution needs to be calculated next to the intercept and offset due to the centring of the variables. Equation 2 shows the general model used to determine the effect of the variables on the outcome.

$$\ln(C) = b_0 + ln(L) + b_{IV} IV + b_M Moderator + b_{int} IV * Moderator \tag{2}$$

4. Results

The results and implications of models containing the previous data and variables are presented in this section. Following the research questions and methodology, and the nature of formed variables, every step required running more than one model. Subsequently, models with base, obstacle, width and light variables and combinations will be discussed. The results of the final models are displayed in Table 4. Green cells indicate significant parameters, while red cells are insignificant parameters with a p-value larger than 0.05 (parameters are found in column B, p-values in column Sig). Dark grey variables refer to parameters that were expected to contribute to bicycle crashes negatively, meaning positive values will increase the numbers. Light grey cells are the other way around. Since interaction effects are not that easy to interpret, they remain white. The Likelihood-Ratio test was used in case of nested models, and the Akaike and Bayesian Information Criterion (AIC/BIC) if not.

First, several models were run with only the base variables: the offset for the road length (not shown in the results, since these have a fixed regression coefficient of 1 and did not have to be estimated) and the motor vehicle and cyclist intensities. This was stated in the methodology and needed as a first check. These are also not shown in the table, but already showed significant positive parameters. This indicates that streets that process more motor vehicles and cyclists generally have higher bicycle crash numbers. This makes sense since there will be more opportunities for crashes. It also corresponds to the earlier study by Wijlhuizen et al. with partially the same data. These stayed significant in all later models. Additions to this base model with variables from Tables 2 and 3 for obstacles, width and light are discussed accordingly.

The first modelling question "What is the relation between obstacles on and around cycling infrastructure and bicycle crashes?" is answered by creating obstacle-crash models. Several variables were alternated due to different levels (obstacle-free zone distance or correlations). The first models did not contain any interactions yet, and since an (interaction) effect for the (visibility of) medians and bollards was only found after combining them that makes distinguishing harder, it was not the final obstacle-crash model OF. From that model additionally follows that a higher share of bottlenecks in the transition from the road surface to the shoulder increases bicycle crash numbers. This makes sense, since cyclists will bump into those first if they move too close to the sides or try to move to the shoulders. A higher share of the closest obstacles in an obstacle-free zone in a distance of 2 metres negatively affects crashes too. This is expected, as when the closest obstacles in a street are nearer, the probability of crashing into them is also higher.

For the third question about the relation between cycling infrastructure width and bicycle crashes, several models with variables related to the width were run. Model WF is the final width-crash model. From this model follows that a higher mean of the width of the dominating facility lowers crash numbers. This makes sense, since a wider stretch to cycle gives more space to avoid potential crashes. The width only became significant when the interactions came into play. The negative two-way track dummy that was also present in the model without interactions indicates that streets with these as the dominating types have lower cyclists crash numbers. This was expected, since streets where cyclists have to cycle mostly on carriageways are least designed for cycling. Interpretation of interactions is also still to be done. These are discussed later.

The model for light conditions with the lamp post share yielded no significant parameter. In the combined model, light will thus be disregarded. For the natural light conditions, cross tables were created for bicycle crashes from 2009 to 2012 (since crashes in the other years did not correspond with the previously used values) at streets with different combinations of obstacle-free zones and the mean of the dominating width, where the date of these crashes is used. It is assumed that crashes in the spring and summer (from March to August) had better light conditions than in the autumn and winter. Set low, average and high values were chosen for the categories. The mean of the obstacle-free zone and width are 0.94 and 1.58 respectively. For the obstacle-free zones, it also is clear that a lot of streets have all of their nearest obstacles in a zone of 2 metres, leading to a share of 1. As low values are not present a lot, the average there was chosen from 0.90 up to 0.95. With 9% and 4.5% of the widths at 1.10 and 2 metres and these values near

Table 5: Observed crash cross table - 2 m

Obstacle-free zone (2 m)	**Mean width**	**Spring & summer**	**Autumn & winter**
Low	Average	47	44
Low	High	50	56
Average	Low	11	7
Average	Average	116	88
Average	High	13	6
High	Low	9	12
High	Average	667	617
High	High	387	371

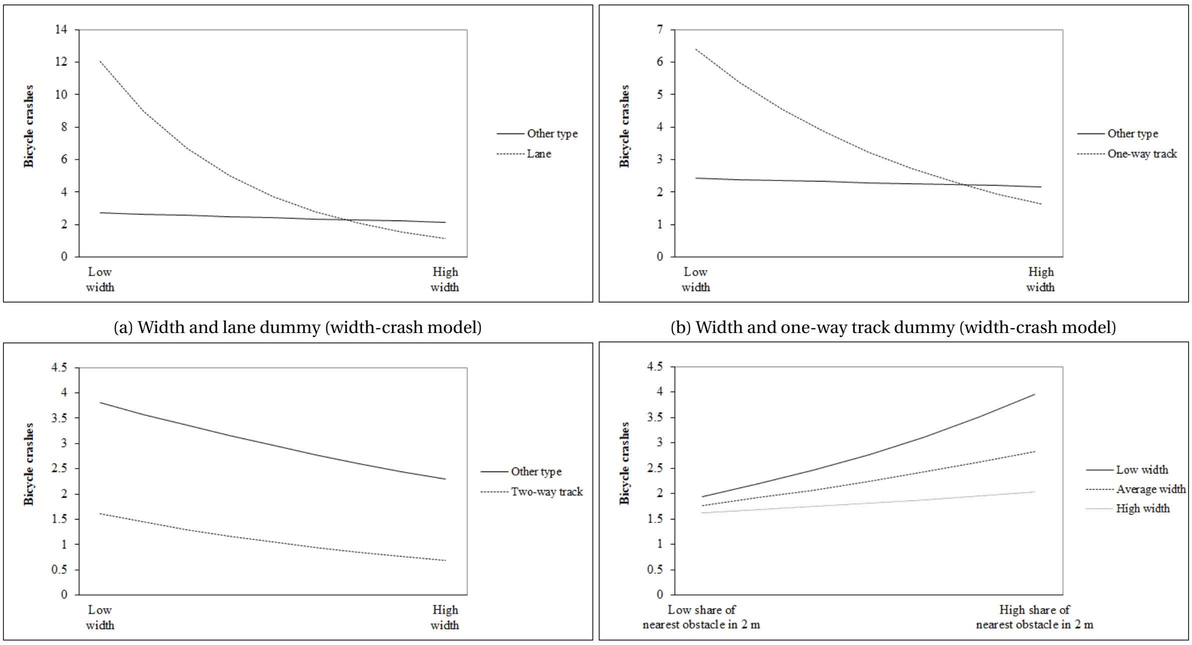
the 25 and 75 percentile cut-offs at 1.14 and 1.95 metres, the average for the width was chosen between these values from 1.10 metre up to 2 metres. The observed crashes are shown in Figure 5. The combination of a low obstacle-free zone share and a low width yielded too low expected values (that were generated by dividing a multiplication of row and column totals by the total crashes in the table), and as such were added to the next category of low obstacle-free zone share and average width. The resulting chi-square value was compared with that in the chi-square table at 95% at the corresponding degrees of freedom ($rows - 1 * columns - 1$, in this case 7). The total of 6.55 is lower than 14.07, meaning that the data cannot indicate a significant difference for bicycle crash numbers in summer and winter. Another cross table was constructed with fewer categories, but that did not point at a significant difference either, even at a lower significance threshold of 90%.

The last modelling question concerned combinations of the previous crash models with obstacle, width and light variables, including interaction terms. For this, the final models were simply combined. Since the transition bottleneck share, obstacle-free zone share and dominating width mean are the significant main effects in the separate models, interaction effects between these were both added (width * transition and width * obstacle-free zone): first with the obstacle-free zone, since its main effect was still significant (not shown), and later together with the transition (Model C2) to test if that would change anything. The most notable differences are the insignificance of the transition bottleneck share and significance of (extra) interaction effects. By combining the predictor variables for different width and obstacle elements, the transition bottleneck share from the data does still not explain bicycle crashes in these combined models. The interaction between the width and obstacle free zone at 2 metres became significant in one model as well. The other main effects stay the same.

Several interaction plots created according to Equation 2 with results from the models in Table 4 are shown in Figure 7. The moderators are depicted by the different lines. The first three plots show the interactions for the width of facility dummies from the width-crash models (almost the same as in the combined model). The widths in the plots vary from one standard deviation from the mean of 1.58 metres at each side, at 95 centimetres and 2.21 metres. The plots indicate that a higher width on lanes and one-way tracks leads to fewer crashes, with the effect reducing at higher widths. This also generally is the case for two-way tracks, but simply due to the main effect of width. The effects of the obstacle-free zone share at 2 metres from the road surface and at three different width levels are shown too. To stick to reasonable values, instead of going up a standard deviation on the horizontal axis at the obstacle-free zone shares at 2 metres, these are stopped at 0.5 standard deviations to the right (about 1). The lines for the interaction effect between the width and the obstacle-free zone share indicate that more obstacles in a 2 metre zone from the road surface lead to more crashes at lower widths. At high widths this has a much lower effect. Lastly, one significant interaction is not shown: visibility problems of medians and bollards. As mentioned, it makes distinguishing harder. However, a

Table 4: Analysis results of several crash models

	Model OF		**Model WF**		**Model C2**		**Model CF**	
	PARAMETERS							
	B	Sig	B	Sig	B	Sig	B	Sig
Intercept	-5.55	0.00	-5.50	0.00	-5.57	0.00	-5.53	0.00
Motor vehicle intensity	0.31	0.00	0.31	0.00	0.32	0.00	0.31	0.00
Cyclist intensity	0.75	0.00	0.71	0.00	0.69	0.00	0.71	0.00
Transition bottleneck share	0.35	0.02			-0.16	0.51		
Obstacle-free zone share - 2 m	2.44	0.00			2.46	0.00	1.75	0.00
Dominating width mean			-0.43	0.01	-0.40	0.01	-0.36	0.02
Lane (L) dummy			0.43	0.03	0.46	0.02	0.39	0.04
One-way track (T1) dummy			0.34	0.10	0.18	0.51	0.22	0.29
Two-way track (T2) dummy			-1.04	0.00	-1.10	0.00	-1.07	0.00
Dominating width mean * L dummy			-1.66	0.01	-1.57	0.01	-1.73	0.00
Dominating width mean * T1 dummy			-0.99	0.00	-0.39	0.37	-0.91	0.00
Dominating width mean * T2 dummy			-0.27	0.63	0.39	0.54	0.19	0.75
Dominating width mean * transition					0.66	0.16		
Dominating width mean * OFZ 2 m					-1.99	0.04		
Negative binomial					0.36		0.38	
	MEASURES OF FIT							
Pearson Chi-Square / df	1.428		1.345		1.289		1.284	
Log-Likelihood	-799.235		-589.214		-584.730		-584.730	
Akaike Information Criterion (AIC)	1610.470		1200.429		1195.460		1193.460	
Bayesian Information Criterion (BIC)	1633.390		1239.888		1242.094		1236.507	



(a) Width and lane dummy (width-crash model)

(b) Width and one-way track dummy (width-crash model)

(c) Width and two-way track dummy (width-crash model)

(d) Width and obstacle-free zone 2 metres (combined model)

Figure 7: Interaction plots

higher problem share led to more crashes.

5. Conclusions

While the Netherlands has a low fatality rate for cyclists, the numbers of severely injured cyclists and cyclist deaths were found to be increasing, cyclists make out a big part of the total injured and deaths, and a reduction in total traffic deaths came to a halt. Together with a research gap in the direction of the obstacle space, that relates to obstacles on the road surface and next to bicycle facilities, and the space (i.e. width of the road) that is available to cycle and in the worst case avoid these obstacles, the importance of cycling in the Dutch culture make it a good research topic. This study therefore looked for an answer to the question *"To what extent can the obstacle space of cycling infrastructure and visibility thereof explain bicycle crashes in Amsterdam?"*. For this, literature research and modelling of generalised linear regression with a negative binomial distribution was used, that used extensive data sources.

From the literature research, it was found that there are many different obstacle types with different functions (such as defining a barrier, controlling access and guarding) and characteristics (on or off the road, fixed or dynamic, and point or line). Crash risk is mainly related to collision, distraction, blocking vision and collision as a result of the latter two. Most common width values from guidelines vary between 1 metre for carriageways to 5 metres for two-directional busy tracks and is dependent on the free-space profile around cyclists and other road users, use of the road and intensities. The available width determines swerving possibilities and, again, the risk to collide with obstacles and other road users.

Through modelling, the next questions were answered. Like in an earlier study by Wijlhuizen et al. (2016) with this data set, firstly it was confirmed that busier streets lead to more crashes (due to more opportunities for crashes), albeit lowered at higher intensities due to the natural logarithm for safety in numbers. Relations between obstacles on and off the road with bicycle crashes were found too. A higher share of nearest obstacles to the road surface and more bottlenecks in the transition from the road surface to shoulder namely result in higher bicycle crashes. Relations between width and bicycle crashes appeared too: wider cycling infrastructure has a positive effect on bicycle crashes. The width of lanes and one-way bicycle tracks even has an additional effect on bicycle crashes. With these results, the hypotheses that more obstacles and narrower infrastructure lead to more crashes were validated. A changing effect between obstacles and crashes was found as well, for an obstacle-free zone at a maximum of 2 metres away from the road surface. The smaller the available space for cyclists is, the higher the crash risk at the same obstacle share. All in all, to a big extent relations between the obstacle space of cycling infrastructure in Amsterdam and registered bicycle crashes in the city exist. While not all variables that were included in the models came out significant, many still did. Apart from visibility problems of on-path obstacles, changing light conditions could not give a reason for crashes in the data set.

6. Discussion

First and foremost, a reassuring finding was that the significant parameters for variables were found to be in the expected direction, thus not giving strange results. Despite this, this study can be improved, especially considering the light conditions. These were now omitted in the combined model, and it is assumed that the effect of obstacles and width is higher in bad light conditions. These namely did reduce safety in earlier studies. All models could be rerun twice, with crashes in well lit and badly lit conditions separated. Furthermore, the effects of the road user and vehicles, that are also part of the conceptual model, were excluded fully in this research due to the focus on infrastructure and the availability of data. Apart from the more general date and time of crashes and in some cases the vehicle type, namely no detailed data about the road user and their vehicle was available, such as age and type of bicycle (especially since older people cycle more during daytime and due to the increase of electric bicycles in the Netherlands). If possible, it would be useful to control for age of cyclists too, since they cycle more during daytime.

The location and data sources also determine the applicability of the results. This study specifically concerned cycling infrastructure along 50 km/h road sections (excluding intersections) inside built-up, metropolitan areas of the municipality of Amsterdam. While the initial set of roads are scattered all through the city, most importantly these do not represent all roads where crashes happened, both in the city itself and the country. It would be useful to also evaluate whether the same relations are present in other cities or countries. Additionally, while some crashes had more detailed locations available that also allow for more detailed analysis of the surroundings that might have led to a crash, all crashes had to be scaled up to the full street level. Albeit rather long, shorter streets in the analyses – 150 metres being the shortest due to filters – result in created variables that are bound to more accurately represent the infrastructure characteristics that were present at the location of the crash.

A further point to make relates to the data quality. Intensities are either taken from an older model or volunteer counts. While the bicycle numbers might better correspond with reality, they certainly do not capture full day averages and all streets. Both intensities on streets are thus merely a proxy of the actual intensity. Nevertheless, they came out significant and predicted crashes well. Road sections can also have changed in the meantime since assessing all characteristics. This was already found to be the case in the process by comparing the actual state of several streets with pictures from 2019 with those used in 2015 and 2016. Despite this, crash data and infrastructure are compared around the same years, and any changes that have happened afterwards are neglected. Related to these points are the registrations. As can be derived from the data set, with 42 per cent of the registered vehicles and 45 per cent of those being cyclists, a big share of the crashes had to be disregarded. Even though registration can be time-consuming at the crash site, underway to the hospital or after treating a crash victim when it is of utmost importance to decrease traffic deaths and injuries, more detailed information is essential when wanting to make more in-depth studies. Parties that compile this data, such as the ambulance and municipalities, should also focus on making the data more complete.

Like the crash data that were predicted at street level due to missing exact locations, infrastructure data were scaled up to higher levels in the form of shares and means. Details on every 25-metre section are most likely disregarded, with only the general impression of that section included. However, it could be that one of these details led to a crash. For the obstacle-free zones, the closest obstacle on the side of the road was registered. This also thus does not allow for determining the clustering of obstacles on a section and does not tell how many are present on that section. For medians and bollards, pretty much the same can be said. Different types of transition and shoulders, and thus obstacles, are not fully explored here yet either. All categories in these would also have split the data too much, leading to few cases. The reduction of usable data already became apparent during the analysis. Combining types of the same form could then be an option.

7. Recommendations

To reduce crash risk, the most straightforward takeaways from this study are of course that cycling infrastructure should come with as little obstacles as possible and be as wide as possible. Problematic medians and bollards on the road regarding visibility lacked profiled road markings and general contrast with the environment. For those obstacles, it is advised to take the guidelines by the CROW into account. The quality of transitions (and shoulders) as assessed depended mainly on the height difference between the road surface and the shoulder. Cracks, bumps and holes are what create these height differences. When these create a high probability of cyclists losing balance and contact should be avoided, they were marked a bottleneck. These bottlenecks should be prevented and repaired. Maintenance is therefore of great essence. Lastly, instead of making cyclists use the carriageway, it is better to make any facility like lanes or tracks available. In the case of lanes, this makes sure at least some dedicated space is available, and for tracks this reduces the exposure to motor vehicles that have to use these facilities. Several carriageways in the assessed streets even did not have enough width available for all road users to drive next to each other. On streets with speed limits at 50 km/h, speed differences are too high and risk increases. These carriageways also still are frequent on streets. From the interaction plots for lanes and one-way tracks even followed that an increase in average width on smaller sections has a bigger impact on crashes than the same increase on wider parts. It is thus a good idea to take a look at improving smaller sections and moving cyclists from the road to more dedicated infrastructure first.

Concluding, this study contributed more empirical evidence to road safety research, that predominantly focused on matters such as the heavier users of infrastructure: motor vehicles. Especially in other countries, cyclist research and developments lag behind, so it is worthwhile for the Netherlands to take the lead. Relations on street level also help focus on anomalies or peculiarities on detailed sections of streets, that could be hot spots for bicycle crashes. The field of cyclist road safety or the more general active mode road safety is not fully explored yet, and more is certainly still to be done.

References

Annema, J., & van Wee, B. (2009). *Verkeer en vervoer in hoofdlijnen (Basics of traffic and transport).* Bussum: Uitgeverij Coutinho.

Bagui, S., & Mehra, K. (2019). On the Convergence of Negative Binomial Distribution. *American Journal of Mathematics and Statistics*, *9*, 44–50. doi: 10.5923/j.ajms.20190901.06

Bax, C. A., van Petegem, J. W. H., & Giesen, M. (2014). *Passen gemeenten de Ontwerpwijzer Fietsverkeer toe?: gebruik van de richtlijnen voor fietsinfrastructuur en factoren die dit beïnvloeden (Do municipalities apply the Design Guide for Bicycle Traffic?: use of the guidelines for cycling infrastructure and influencing factors).* Den Haag, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.

Bijleveld, C. C. J. H., & Commandeur, J. J. F. (2012). *Multivariate analyse: een inleiding voor criminologen en andere sociale wetenschappers (Multivariate analysis: an introduction for criminologists and other social scientists).* Den Haag, Boom Lemma Uitgevers, 2012, XXII+ 417 p., ref.; 3e gew. druk.

Buehler, R., & Pucher, J. (2012). Walking and cycling in Western Europe and the United States: trends, policies, and lessons. *Tr News*(280).

Castro, A., Kahlmeier, S., & Gotsch, T. (2018). Exposure-adjusted road fatality rates for cycling and walking in European countries..

Copenhagenize Design Company. (2020). *2019 Copenhagenize Index.* Retrieved from `https://copenhagenizeindex.eu/`

CROW. (2018a). *Fietsverkeer: Ontwerpwijzer bruggen voor langzaam verkeer (Bicycle traffic: Design guide for bridges for slow traffic).* Retrieved 2020-02-05, from `https://kennisbank.crow.nl/Kennismodule`

CROW. (2018b). *Fietsverkeer: Ontwerpwijzer fietsverkeer (Bicycle traffic: Design guide for bicycle traffic).* Retrieved 2020-02-05, from `https://kennisbank.crow.nl/Kennismodule`

Dawson, J. (2014, 03). Moderation in management research: What, why, when, and how. *Journal of Business and Psychology*, *29.* doi: 10.1007/s10869-013-9308-7

DiGioia, J., Watkins, K. E., Xu, Y., Rodgers, M., & Guensler, R. (2017). Safety impacts of bicycle infrastructure: A critical review. *Journal of safety research*, *61*, 105–119.

Dziak, J. J., & Henry, K. L. (2017). Two-part predictors in regression models. *Multivariate behavioral research*, *52*(5), 551–561.

Fabriek, E., De Waard, D., & Schepers, J. P. (2012). Improving the visibility of bicycle infrastructure. *International journal of human factors and ergonomics*, *1*(1), 98–115.

Gemeente Amsterdam. (2016, September). *Meerjarenplan Verkeersveiligheid 2016-2021 (Multi-year Road Safety Plan 2016-2021).*

Handy, S., Van Wee, B., & Kroesen, M. (2014). Promoting cycling for transport: research needs and challenges. *Transport reviews*, *34*(1), 4–24.

Holsclaw, T., Hallgren, K. A., Steyvers, M., Smyth, P., & Atkins, D. C. (2015). Measurement error and outcome distributions: Methodological issues in regression analyses of behavioral coding data. *Psychology of addictive behaviors*, *29*(4), 1031.

Hutchinson, M. K., & Holtman, M. C. (2005). Analysis of count data using Poisson regression. *Research in Nursing & Health*, *28*(5), 408–418.

Kennisnetwerk SPV. (2018). *Veilig van deur tot deur: het Strategisch Plan Verkeersveiligheid 2030 (Safely from door to door: the Strategic Plan for Road Safety 2030).*

Klok Real Estate. (n.d.). *Stadsdelen Amsterdam (City districts of Amsterdam).* Retrieved 2020-02-19, from `https://www.klokrealestate.nl/124-8bcc/klok-real-estate-informatie-stadsdelen-amsterdam`

McCullagh, P., & Nelder, J. (1989). *Generalized linear models second edition chapman & hall.* London.

Ormel, W., Klein Wolt, K., & Den Hertog, P. (2008). Enkelvoudige fietsongevallen (Single-bicycle crashes). *Amsterdam: Stichting Consument en Veiligheid.*

Othman, S., Thomson, R., & Lannér, G. (2009). Identifying critical road geometry parameters affecting crash rate and crash type. In *Annals of advances in automotive medicine/annual scientific conference* (Vol. 53, p. 155).

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & quantity*, *41*(5), 673–690.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *50*(302), 157–175.

Reynolds, C. C. O., Harris, M. A., Teschke, K., Cripton, P. A., & Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. *Environmental health*, *8*(1), 47.

Schaap, T. W., Harms, L. W. J., Kansen, M., & Wüst, H. (2015). *Fietsen en lopen: de smeerolie van onze mobiliteit (Cycling and walking: the lubricant of our mobility).* Kennisinstituut voor Mobiliteitsbeleid.

Schepers, P. (2008). De rol van infrastructuur bij enkelvoudige fietsongevallen (The role of infrastructure in single-bicycle crashes). *Directoraat-Generaal Rijkswaterstaat, Dienst Verkeer en Scheepvaart DVS, Delft*.

Schepers, P., & den Brinker, B. (2011). What do cyclists need to see to avoid single-bicycle crashes? *Ergonomics*, *54*(4), 315–327.

Schepers, P., Hagenzieker, M., Methorst, R., Van Wee, B., & Wegman, F. (2014). A conceptual framework for road safety and mobility applied to cycling safety. *Accident Analysis & Prevention*, *62*, 331–340.

Schepers, P., & Klein Wolt, K. (2012). Single-bicycle crash types and characteristics. *Cycling Research International*, *2*(1), 119–135.

Schepers, P., Stipdonk, H., Methorst, R., & Olivier, J. (2017). Bicycle fatalities: trends in crashes with and without motor vehicles in The Netherlands. *Transportation research part F: traffic psychology and behaviour*, *46*, 491–499.

Stipdonk, H. L., & Bos, N. (2014). *Verkeersveiligheidsanalyse gemeente Amsterdam: relevante doelgroepen, ontwikkelingen en aandachtspunten (Road safety analysis for the municipality of Amsterdam: relevant target groups, developments and points of attention).* Den Haag, SWOV.

SWOV. (2012). *Risico in het verkeer (Risks in traffic).* (SWOV factsheet)

VeiligheidNL. (2019). *Aantal fietsers met ernstig letsel op de Spoedeisende Hulp gestegen (Number of cyclists with serious injuries to the emergency room increased).* Retrieved 2019-12-05, from `https://www.veiligheid.nl/verkeer/actueel/aantal-fietsers-met-ernstig-letsel-op-de-spoedeisende-hulp-gestegen`

Wanvik, P. O. (2009). Effects of road lighting: an analysis based on Dutch accident statistics 1987–2006. *Accident Analysis & Prevention*, *41*(1), 123–128.

Wardlaw, M. J. (2014). History, risk, infrastructure: perspectives on bicycling in the Netherlands and the UK. *Journal of Transport & Health*, *1*(4), 243–250.

Weijermars, W. (2019). *Monitor Verkeersveiligheid 2019: effectieve maatregelen nodig om het tij te keren (Road Safety Monitor 2019: effective measures are required to stem the tide).* Den Haag, SWOV.

Weijermars, W., Van Schagen, I., Aarts, L., Van Petegem, J., Van Stipdonk, H., & Wijnen, W. (2018). *Hoe verkeersveilig kan Nederland zijn in 2030?: mogelijkheden voor reductie in aantallen verkeersslachtoffers (How safe can the Netherlands be in 2030?: possibilities for traffic victim reductions).* Den Haag, SWOV.

Wijlhuizen, G. J., van Petegem, J. W. H., Goldenbeld, C., van Gent, P., de Bruin, J., de Commandeur, J. J. F., & Kars, V. (2016). *Doorontwikkeling CycleRAP-instrument voor veiligheidsbeoordeling fietsinfrastructuur: doelmatigheid handmatige intensiteitsmetingen, betrouwbaarheid beoordelingen infrastructuur en validiteit van het CycleRAP-instrument (Further development of the CycleRAP instrument for safety assessment of cycling infrastructure: effectiveness of manual intensity measurements, reliability of infrastructure assessments and validity of the CycleRAP instrument)).* Den Haag, SWOV.

Wijlhuizen, G. J., van Petegem, J. W. H., Schermers, G., de Bruin, J., de Commandeur, J. J. F., & Kars, V. (2017). *Ontwikkeling Netwerk Safety Index gemeente Amsterdam (Network Safety Index development for the municipality of Amsterdam).* Den Haag, SWOV.

Wood, A. G., Mountain, L. J., Connors, R. D., Maher, M. J., & Ropkins, K. (2013). Updating outdated predictive accident models. *Accident Analysis & Prevention*, *55*, 54–66.

Yan, J., Guszcza, J., Flynn, M., & Wu, C.-S. P. (2009). Applications of the offset in property-casualty predictive modeling. In *Casualty actuarial society e-forum, winter 2009* (p. 366).

B

SWOV

The Dutch national institute for road safety research SWOV is an independent research institute in the Netherlands that was founded in 1962. Its goal is to improve road safety. To accomplish this, it gathers, uses and distributes knowledge. It has ten research themes across four research departments: Infrastructure and Traffic, Road User Behaviour, Human Factors and Vehicle Automation, and Data and Analysis for Policy. For the most part, its research is laid out in a programme arranged with the Ministry of Infrastructure and Water Management. In the past, among others, it conducted research for governmental bodies such as provinces and local governments, the Dutch Driving Test Organisation and an insurance company, and worked together with the Dutch cyclists' union (SWOV, n.d.). It started the vision of "Duurzaam Veilig Wegverkeer" ("Sustainable Safety"), that strives for maximum road safety for everyone. The current third version re-evaluates if the one that was created in the nineties suffices to last to 2030 (SWOV, 2018). Examples of recent research published in 2019 are "Determinants and barriers of walking, cycling and using Personal e-Transporters", "A taxonomy of potential impacts of connected and automated vehicles at different levels of implementation" and "Severe traffic injuries with taxis" (translated) (SWOV, 2019).

SWOV
INSTITUTE FOR
ROAD SAFETY RESEARCH

C

Data assessment and contents

On the following pages, instructions are presented that were used for assessing the cycling infrastructure in Amsterdam, and thus an overview of the data that can be found in the data set (Wijlhuizen et al., 2017). All options regarding this assessment as present in the data set are found in Table C.1 thereafter.

C.1. Instructions for cycling infrastructure assessment (in Dutch)

See next page.

Safe Cycling Network

Handleiding voor beoordelaars van kenmerken van veiligheid van fietsinfrastructuur

Oktober 2016

Kenmerken van de te beoordelen fietsinfrastructuur;

Kijk bij beoordeling ook op plattegrond links in het scherm

Kruispunt in 1 profiel scoren; ook bij groot kruispuntvlak

Soort Kruising		**Fietsers kruisen een GOW, ETW.**	**NB.: Kruising fietsvoorzieningen beschouwen als uitrit**
1	Geen		
2	Gelijkwaardig	Rechts heeft voorrang	
3	Voorrangsregeling	Haaietanden, voorrangsbord	Een drempel is geen voorrangsregel; een **uitritconstructie*** wel
4	VRI	Verkeersregel installatie (verkeerslichten)	
5	Rotonde	Punaise is geen rotonde	Kan ook met verkeerslichten zijn geregeld; dan als VRI beoordelen
6	Oversteek	**Als de fietsvoorziening >10 meter afligt van een kruispunt.**	**Kan ook met verkeerslichten zijn geregeld; dan als VRI beoordelen**
1	Drietaks	**Fietspad zonder rijbaan dat aansluit**	
2	Viertaks		
3	>Viertaks		
4	**Bajonet**		
5	**Anders**		
6	**NVT**		

***Voorbeeld van een uitritconstructie als voorrangsregeling**

Soort Kruising

2. Gelijkwaardig kruispunt

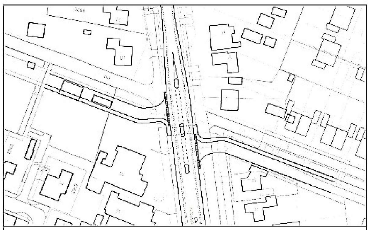
3. Voorrangsregeling

4. Verkeersregel installatie (VRI)

5. Rotonde

6. Oversteek

Bajonet kruising (2 x Drietaks)



Zicht kruispunt (heen en terug)

1	Geen belemmering van zicht	
2	Enigszins belemmerd zicht	Wel kijken (links of rechts) maar doorrijden zonder in te hoeven houden.
3	Ernstig belemmerd	Gericht (links of rechts) kijken, voor voldoende zicht en veiligheid moeten afremmen/inhouden.
4	NTB	

2 Enigszins belemmerd

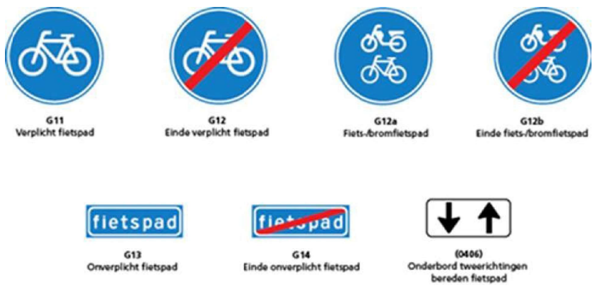
3 Ernstig belemmerd

Fietsvoorziening

1	(Brom) fietspad	Fietspad, aangeduid door bord G12a (gewijzigd RVV 1990), en toegestaan voor fietsers, snorfietsers en bromfietsers. (toegepast in situaties waarin het ongewenst is dat bromfietsers gebruikmaken van de rijbaan voor snelverkeer);
2	Fietspad	ad bestemd voor fietsers, snorfietsers.
3	Fietsstrook	Strook **met** fietsafbeelding op wegdek
4	Fietssuggestiestrook	Strook **zonder** fietsafbeelding
5	Rijbaan	Geen markering voor fietser
6	Anders, namelijk:	Bijvoorbeeld fietsstraat
7	NTB	Niet te bepalen

Om de status van fietsvoorzieningen te kunnen onderscheiden, wordt onderstaande bebording op de openbare weg toegepast (*Afbeelding 3.1*).

- G11 verplicht fietspad, G12 einde verplicht fietspad
- G12a verplicht fiets-/bromfietspad, G12b eind verplicht fiets-/bromfietspad
- G13 onverplicht fietspad (verboden voor brom- en snorfietsen met motor in werking), G14 einde onverplicht fietspad



Bebording die de juridische status aangeeft van fietsvoorzieningen

Fietsvoorziening

1. (Brom)fietspad:

2. Fietspad:

3. <u>Fietsstrook:</u> Verharding fietsstrook met fietssymbool in rode kleur.

4. <u>Fietssuggestiestrook:</u> Verharding suggestiestrook zonder fietssymbool. Met onderbroken markering.

5. <u>Rijbaan:</u> Geen rijbaanindeling; geen scheiding van verkeerssoorten

6. <u>Fietsstraat:</u> Hoogwaardige fietsverbinding, met medegebruik door gemotoriseerd verkeer.

Bijzondere voorziening fiets		
1	Tunnel	
2	Brug	
3	Fly-over	
4	Anders, namelijk:	
5	NVT	
6	***NTB***	

Rijrichtingen fietsvoorziening		
1	Eenrichtingspad	Fietser op de rijbaan zonder strook: 2 keer (heen, terug) als 1 richting beoordelen
2	Tweerichtingspad	
3	***NTB {Niet te bepalen}***	
Zicht rijbaan rechtsafslaan		
1	Geen belemmering van zicht	In welke mate kunnen afslaande auto's fietsers zien (bijvoorbeeld vanwege geparkeerde auto's en andere obstakels)
2	Enigszins belemmerd zicht	
3	Ernstig belemmerd	
4	NVT	
5	***NTB***	
Rijrichting rijbaan		Rijrichtingen zijn te bepalen aan de hand van borden die aan het begin en bij elke kruising van de weg staan.
1	Eenrichting	
2	Tweerichtingen	
3	NVT	
4	***NTB***	

Ligging		
1	Vrijliggend, Binnen bebouwde kom	Apart van rijbaan, gescheiden door fysieke voorziening (bijvoorbeeld: niveauverschil/ beplanting). Ook fietspad door park niet langs rijbaan; dat zijn de solitaire fietspaden.
2	Vrijliggend, Buiten bebouwde kom	
3	Aanliggend Binnen bebouwde kom	
4	Aanliggend, Buiten bebouwde kom	
5	NVT / onbekend	Situaties waarbij geen sprake is van een fietspad.
6	NTB	

1. Vrijliggend fietspad
   Fietspad dat hetzij parallel loopt met de naastgelegen rijbaan en daarvan door een tussenberm wordt gescheiden, hetzij een geheel eigen tracé volgt.

2. Aanliggend fietspad
   Fietspad dat door een zeer smalle tussenberm is gescheiden van de naastgelegen rijbaan.

Omgeving

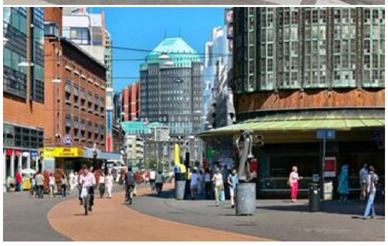
1	Recreatiegebied, park, bos	Grote kans medegebruik
2	Winkel-/ uitgaansgebied/ scholen	Grote kans medegebruik
3	Anders, namelijk:	Grote kans medegebruik
4	Geen medegebruik (o.a.: wonen, bedrijven)	
5	*NTB*	

Medegebruik fietsvoorziening

Fietsvoorziening die door winkelgebieden of recreatiegebieden loopt waardoor kans op medegebruik door bijvoorbeeld wandelaars of hardlopers/ skeelers groot is. We onderscheiden twee categorieën waar de kans op medegebruik groot is:

- fietsvoorzieningen door recreatiegebieden
- fietsvoorzieningen door winkelgebieden/ uitgaansgebied/ scholen.

In deze gebieden is de kans op botsingen relatief groot, omdat de medegebruikers zich met andere snelheden en doelen (bijvoorbeeld winkelen, hardlopen of skeeleren) op de fietsvoorziening bevinden.



Verhardingsbreedte fietsvoorziening

- Bij tweerichtingen fietspad de volledige breedte bepalen
- Bepalen met behulp van meet-tool in cyclorama
- Bij fietsers op de rijbaan (zonder fietssuggestiestrook) standaard 1,5 meter per richting.

Verharding-Type fietsvoorziening (ook de rijbaan beschouwen als de fietsers daar moeten fietsen)

1. **Open** **(klinkers, tegels, keien, grind)**
2. **Gesloten** **(asfalt, beton)**
3. **NTB**

Verharding Kwaliteit
De fiets is een balansvoertuig dat door oneffenheden uit evenwicht kan raken, waardoor zwenkingen optreden om het evenwicht te herstellen of de fietser ten val komt. Bij de kwaliteit van de verharding is de aandacht dan ook gericht op de aanwezigheid van oneffenheden in de verharding:

- Scheuren (S),
- Gaten (G)
- Hobbels (H) – SGH.

Het gaat daarbij ook om de rand van de verharding naar de berm. Deze kan zijn afgebrokkeld of gescheurd door bijvoorbeeld verzakking van de berm. Bij de beoordeling gaat het ook om aangebrachte voorzieningen die oneffenheden kunnen veroorzaken, zoals putten, tramrails of wildroosters.

NB.: Als fietsers op de rijbaan fietsen (zonder eigen ruimte), dan een strook van naar schatting <u>1,5 meter vanaf de berm</u> beoordelen!

1	Voldoende	Vrijwel geen sprake van S,G,H.
2	Aandachtpunt	Geringe mate van S,G,H; geen acuut gevaar voor uit balans raken; wel oncomfortabel (bijv.: stalen putdeksel in verharding)
3	Knelpunt	Grote mate van S,G,H, grote kans op uit balans raken; contact vermijden
4	NTB	

2. <u>Aandachtspunt</u>

3. <u>Knelpunt</u>

Uitritten
1. Ja
2. Nee

Is er tenminste 1 uitrit per 25 meter straat of fietspad. Het gaat om een uitrit voor voertuigen van een gebouw of perceel naar de openbare weg, en/of de ingang voor voertuigen vanaf de openbare weg. Ook aansluitingen/ kruisingen van fietspaden onderling. Voorbeelden zijn:

- de oprit naar een garage of carport;
- de oprijlaan bij een landhuis of landgoed;
- de toegang tot een weiland of bosperceel;
- de ingang van een parkeergarage;
- de toegang tot een bedrijventerrein;
- aansluitend/ kruisend fietspad.

NB.: Parkeervakken op of langs de rijbaan worden niet als uitrit beschouwd; deze worden apart gescoord.

Voorbeeld: aansluitend/ kruisend fietspad

Bocht scherp

1	Geen bocht of onscherpe bocht	Kunnen doortrappen bij bocht nemen zonder op andere helft te komen.
2	Scherpe bocht.	Bij nemen van bocht snelheid omlaag, trappers stilhouden, grote kans om op andere helft te komen.
3	NTB	

1. Geen bocht of onscherpe bocht

2. Scherpe bocht (naar links)

Bocht Zicht

1	Geen belemmering van zicht	
2	Enigszins belemmerd zicht	Wel kijken maar doorfietsen zonder in te hoeven houden.
3	Ernstig belemmerd	Gericht kijken, voor voldoende zicht en veiligheid moeten afremmen/inhouden.
4	NTB	

2. Enigszins belemmerd zicht

3. Ernstig belemmerd zicht

Snelheidslimiet rijbaan

De snelheidslimiet op de rijbaan is van belang op plaatsen waar fietsers het verkeer op de rijbaan kruisen of samen gebruikmaken van dezelfde rijbaan. Relatief grote snelheidsverschillen bepalen in belangrijke mate de ernst van het letsel van een fietser bij een aanrijding met bijvoorbeeld een auto.

De volgende hoofdcategorieën van wegen worden onderscheiden, met de daaraan gerelateerde geldende maximumsnelheden:

- *Stroomwegen (SW):* wegen met een primaire verkeersfunctie, bedoeld voor een zo veel mogelijk conflictvrije afwikkeling van gemotoriseerd verkeer. Stroomwegen kenmerken zich door een fysieke rijbaanscheiding en ongelijkvloerse kruisingen. In Nederland kunnen fietsers niet kruisen met een stroomweg.

- *Gebiedsontsluitingswegen (GOW):* wegen die zowel doorstroming als uitwisselen tot doel hebben. Gebiedsontsluitingswegen kenmerken zich door scheiding van snel- en langzaam verkeer (parallelle fietspaden) en gelijkvloerse kruisingen. Buiten de bebouwde kom mag er door snelverkeer 80 km/uur gereden worden, binnen de bebouwde kom 50 km/uur of 70 km/uur.

- *Erftoegangswegen (ETW):* wegen met een verblijfsfunctie, bestemd om percelen toegankelijk te maken. Erftoegangswegen hebben geen rijbaanscheiding en snel- en langzaam verkeer rijdt gemengd (mogelijk ook medegebruik), hetgeen een relatief lage maximumsnelheid vereist. Doorgaand verkeer wordt bij voorkeur zo veel mogelijk geweerd. Buiten de bebouwde kom mag op erftoegangswegen 60 km/uur gereden worden, binnen de bebouwde kom 30 km/uur.

1 15 km
2 30 km
3 50
4 60
5 70
6 80
7 100
8 120
9 130

Versmalling

1	Geen of nauwelijks	Vrijwel geen verandering in koers nodig, bijvoorbeeld bij geleidelijke versmalling
2	Aanmerkelijk	Actief koers aanpassen door sturen. Met name als versmalling plotseling is (oa. door object). Bijvoorbeeld: fietsers naast elkaar moeten achter elkaar gaan rijden**. Let op: dit is ook het geval bij auto's die op de rijbaan mogen parkeren als fietsers op de rijbaan rijden**
3	NTB	

2. <u>Aanmerkelijke versmalling</u>

Hoogteprofiel

1	Vlak	
2	Stijging/daling	Lagere trapfrequentie, meer kracht of fietsen in lagere versnelling bij oprijden van helling.
3	NTB	

2. Stijging/daling

Tramrails

1 Nee

2 Ja, in gedeelde ruimte (van fietsers of auto’s)

3 Ja, in gescheiden ruimte (van fietsers of auto’s)

4 NTB

Overgang Kwaliteit

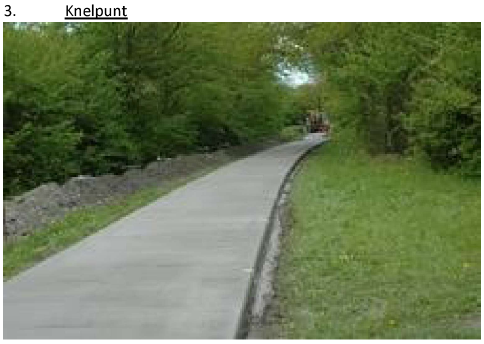
*Vooral de mate van hoogteverschil tussen het fietspad en de berm is van belang. Een berm kan bijvoorbeeld zijn kapotgereden waardoor direct naast het fietspad een kuil is. Ook kan bijvoorbeeld het asfalt van het fietspad hoger liggen dan de berm. Soms is er geen overduidelijke overgang, zoals bij aaneengesloten struiken/ hek dat **direct aansluit** op het fietspad of bij een kade. Dat zijn knelpunten; er is geen ruimte om uit te wijken naar een berm je valt of botst dan onmiddellijk*

1	Voldoende (vlak)	
2	Aandachtspunt	Geringe mate van hoogteverschil; S,G,H; geen acuut gevaar voor uit balans raken; wel oncomfortabel
3	Knelpunt	Grote mate van hoogteverschil; S,G,H, grote kans op uit balans raken; contact vermijden **(bijv. stoepranden, maar ook geparkeerde auto's)**
4	NTB	

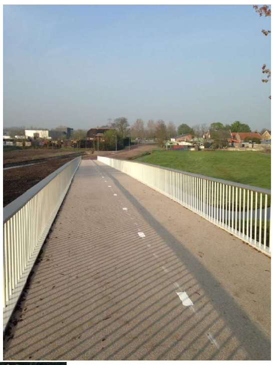
1. <u>Voldoende</u>

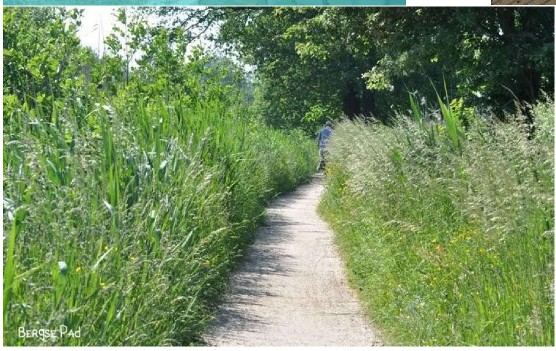

2. <u>Aandachtspunt</u>

3. <u>Knelpunt</u>










Overgang Type

1	Vlak	
2	Opsluitband - overrijdbaar	Sterk afgevlakt oplopende opsluitband (kan zijlings worden opgereden door fiets zonder balansverstoring)
3	Opsluitband - niet overrijdbaar	Schuin oplopende of rechthoekige afsluitband die balansverstoring geeft als die zijdelings wordt aangereden door fiets.
4	Scherpe rand wegdek (beton / stelcon)	
5	Geul	Bijvoorbeeld voor afvoer van regenwater naar kolk
6	Hek, hoge opstaande rand	Direct aansluitend aan de verharding
7	Parkeervak op rijbaan	**Ook als er op het plaatje geen auto staat**
8	Parkeervak langs rijbaan	**Ook als er op het plaatje geen auto staat**
9	Anders, namelijk:	Bijvoorbeeld begroeiing of kade
10	NTB	

Straatverlichting

Hoeft niet elke 25 meter; als er wel regelmatig verlichting is aangebracht dan is er verlichting

1 Aanwezig
2 Niet aanwezig
3 NTB

Markering

1 Links — Gezien vanuit de fietser links
2 Rechts — Lijn aan berm-kant van fietsvoorziening **(voor fietser rechts; bij 2 richtingspad minimaal 2 lijnen)**
3 Links en Rechts
4 Links en rechts en midden — In geval van tweerichtingsfietspad
4 Geen markering
5 NTB

Paal in pad

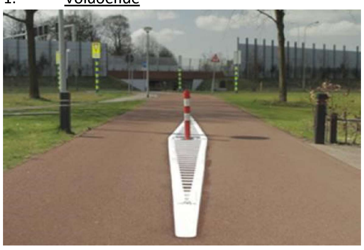
Het gaat om een paal in de verharding zelf. Als de paal op een verhoging staat (eiland) dan is het een obstakel in de berm.

1 Ja
2 Nee
3 NTB

Paal Zicht

1 Voldoende — Goed zichtbaar paaltje (verlicht en contrasterend met achtergrond) met ribbelmarkering
2 Aandachtspunt — Geen ribbelmarkering/ wel goed zichtbaar
3 Knelpunt — Geen ribbelmarkering/ niet goed zichtbaar
4 NVT
5 NTB

1. <u>Voldoende</u>



2. <u>Aandachtspunt</u>

3. <u>Knelpunt</u>

Het gaat om een verhoogd obstakel dat aan de linkerkant van de fietser is geplaatst als geleiding en/of versmalling.

Middeneiland aanwezig

Een paal op een eiland is een obstakel in de berm en niet een paal in pad; zie voorbeeld.

1	Ja
2	Nee
3	NTB

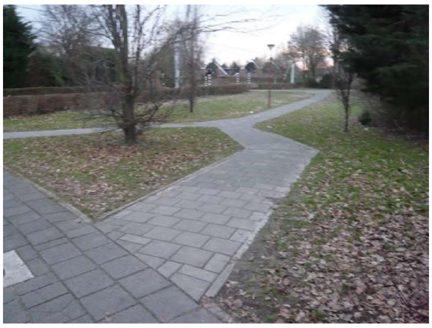


Middeneiland Zicht

1	Voldoende	Goed zichtbaar middeneiland (verlicht en contrasterend met achtergrond) met ribbelmarkering
2	Aandachtspunt	Geen ribbelmarkering/ wel goed zichtbaar (verlicht en contrasterend met achtergrond)
3	Knelpunt	Geen ribbelmarkering/ niet goed zichtbaar
4	NVT	
5	NTB	

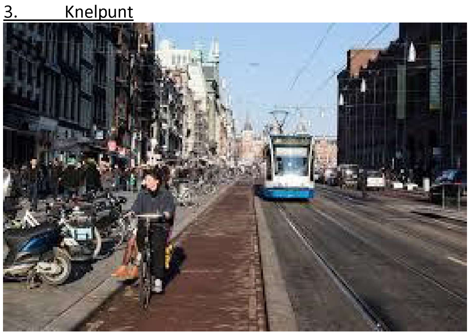
Berm Kwaliteit		**Let op: Afgezien van de overgang naar de berm (stoeprand kan knelpunt zijn maar de stoep zelf voldoende)**
1	Voldoende	Goed om op te fietsen: vlak en zonder obstakels binnen 1 meter
2	Aandachtspunt	Geen acuut gevaar voor uit balans raken; wel oncomfortabel
3	Knelpunt	Grote kans op uit balans raken (S,G, H); contact vermijden (ook bij in berm geplaatst groen zoals: heg, struik, of muur, hekwerk, talud, etc ***binnen 1 meter van de fietsverharding***). NB.: Let op afstand naar rijbaan. Als die binnen 1 meter is, dan ook knelpunt scoren ivm kans op aanrijding met motorvoertuigen.
4	NTB	

1. Voldoende

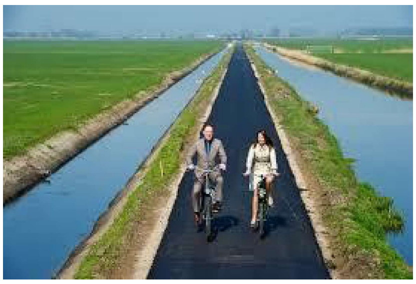


2. Aandachtspunt

3. Knelpunt



3. Knelpunt



Berm type **Bij combinaties (bijvoorbeeld eerst gras dan een talud en dan een sloot) het eerste risico 'het talud' aangeven.**

1 Gras
2 Aaneengesloten begroeiing met planten, struiken, heg in de berm ***binnen 1 meter***
3 Aarde/ zand / klei
4 Steenslag/ grind
5 Verharding; bijvoorbeeld trottoir
6 Parkeervak/ gelegenheid ***binnen 1 meter***
7 Sloot/ kanaal ***binnen 1 meter***
8 Talud ***binnen 1 meter***
9 Berm is geblokkeerd aansluitend aan fietsvoorziening
10 Hek of muur ***binnen 1 meter***
11 Anders, namelijk:
12 NTB

Obstakel afstand

Het gaat om obstakels **(links of rechts**) waartegen gebotst kan worden (palen, bomen, geparkeerde voertuigen etc.). Ook kunnen het andere gevaren zijn als vallen van een kade of in sloot/kanaal. ***Obstakel met kortste afstand beoordelen.***

1 Aangrenzend aan verharding
2 <0.5 meter
3 0.5-1 meter
4 1-2 meter
5 >2 meter
6 NVT

Werk in uitvoering

1 Ja
2 Nee
3 NTB

C.2. Input for cycling infrastructure assessment

Table C.1: Input for assessment of bicycle facility characteristics (translated from Wijlhuizen et al. (2017)

	Characteristic	Input	Used
1	Type of crossing	None Level Priority Traffic signal control Roundabout Crossing	
2	Number of arms	Three Four More than four Bayonet intersection Other Not applicable	
3	Vision on intersection	No obstructed view Slightly obstructed view Significantly obstructed view Could not be determined	
4	Type of bicycle facility	Bicycle-moped track Bicycle track Bicycle lane Non-designated bicycle lane Carriageway Other Not applicable / Could not be determined	Yes
5	Special bicycle facility	Tunnel Bridge Fly-over Other Not applicable / Could not be determined	
6	Number of directions bicycle track	One-way Two-way Could not be determined	Yes*
7	Vision on intersection (cyclist to the right)	No obstructed view Slightly obstructed view Significantly obstructed view Not applicable / Could not be determined	
8	Number of directions carriageway	One-way Two-way Not applicable / Could not be determined	
9	Location of bicycle track	Solitary, inside built-up area Solitary, outside built-up area Non-solitary, inside built-up area Non-solitary, outside built-up area Not applicable / Could not be determined	
10	Surroundings	Recreational area, park, forest Shopping area, nightlife area, schools Other, namely... No shared use (living, companies) Could not be determined	
11	Pavement width	In centimetres; full width for two-way bicycle tracks	Yes

Table C.1 continued from previous page

	Characteristic	Input	Used
12	Pavement type	Open Closed Could not be determined	
13	Pavement quality	Sufficient Point of attention Bottleneck Could not be determined	
14	Exits present (per 25 metres)	Yes No	
15	Sharp turn	No turn or mild turn Sharp turn Could not be determined	
16	Vision on turn	No obstructed view Slightly obstructed view Significantly obstructed view Not applicable / Could not be determined	
17	Speed limit carriageway	15, 30, 50, 60, 70, 80, 100 km/h	Yes*
18	Narrowing	None or barely Significantly Could not be determined	Yes
19	Height profile	Flat Upward/downward slope Could not be determined	
20	Tram rails	No Yes, in shared space (of cyclists or motor vehicles) Yes, in separated space (of cyclists or cars) Could not be determined	
21	Transition quality	Sufficient (flat) Point of attention Bottleneck Could not be determined	Yes
22	Transition type	Flat Kerb - passable Kerb - impassable Sharp edge of road surface (concrete/stelcon) Trench Fence, raised edge Parking on carriageway Parking along carriageway Other Could not be determined	No**
23	Street lighting	Present Not present Could not be determined	Yes
24	Markings	Left Right Left and right No markings Could not be determined	
25	Bollard on road	Yes No Could not be determined	Yes

Table C.1 continued from previous page

	Characteristic	Input	Used
26	Visibility of bollard	Sufficient Point of attention Bottleneck Not applicable / Could not be determined	Yes
27	Median on road	Yes No Could not be determined	Yes
28	Visibility of median	Sufficient Point of attention Bottleneck Not applicable / Could not be determined	Yes
29	Shoulder quality	Sufficient Point of attention Bottleneck Not applicable / Could not be determined	Yes
30	Shoulder type	Grass Continuous vegetation within 1 metre Soil/sand/clay Crushed stone/gravel Road surface (e.g. pavement) Parking place Ditch/channel within 1 metre Slope Shoulder is blocked adjacent to facility Fence or wall within 1 metre Other Could not be determined	No**
31	Obstacle-free zone next to bicycle facility	Adjacent to pavement <0.5 metre 0.5 - 1 metre 1 - 2 metres >2 metres Not applicable	Yes
32	Roadworks	Yes No Not applicable / Could not be determined	

* For data filtering and preparing
** Can be considered for later crash models

C.3. Input for motor vehicle infrastructure assessment

Table C.2: Input for assessment of motor vehicle facility characteristics (translated from Wijlhuizen et al. (2017))

	Characteristic	Input	Used
1	Location of road section	Road section Intersection Near intersection	
2	Road category (Sustainable Safety)	Access roads Distributor roads Through-roads	
3	Obstacle-free zone (in cm)	Number	
4	Public transport stops	None On carriageway Separate lane/carriageway (left, right, middle) Lay-by	
5	Break-down facility	No break-down facility Emergency lane Break-down bay Shoulder	
6	Advance notice signposting	Present Absent	
7	Directional separation	None (examples in and outside built-up area) Impassable Hardly possible Single barrier line Double barrier line Single broken line Double broken line Double filled centre line marking Passable tram/bus lane Impassable tram/bus lane Nose Other	Yes*
8	Edge marking or facility	None Barrier line Broken line	
9	Access restriction	None (all road users allowed) Pedestrians Bicycle Moped Bicycle/moped Slow traffic (fully restricted) Other (including combinations of the above or e.g only restricted for freight traffic; width restriction; motor vehicles with more than four wheels, etc.)	
10	Road surface	Open Closed	
11	Speed limit	15; 30; 50; 60; 70; 80; 90; 100; 120; 130 km/h or 999 if unknown	
12	Parallel road	None Service road Unknown Could not be determined	

Table C.2 continued from previous page

	Characteristic	Input	Used
13	Parking	None Allowed (no markings) Parking on the road Parking next to the road Not allowed	
14	Exits (access)	Number per 25 m	
15	Speed hump/raised section (physical speed reduction measure)	None Speed hump Median Offset Other	
16	Bicycle/moped facility on carriageway	None Bicycle lane Mandatory bicycle track Mandatory bicycle-moped track Non-mandatory bicycle-moped track Unknown bicycle(-moped) track (no sign present) (Adjacent) non-dedicated bicycle lane (no image of bicycle) Adjacent bicycle lane (with image of bicycle)	
17	Width of rescue lane	Number	
18	Straight section or curve (length)	Yes, straight section No, no straight section	
19	Intersection type	None Bicycle(-moped) crossing At-grade (no roundabout) At-grade with raised section Roundabout Grade separated (viaducts, etc.) Combination	
20	Road width (in cm)	Number	Yes*
21	Number of lanes	Number	Yes*
22	Road view	Open (in/outside built-up area) Half (in/outside built-up area) Closed (in/outside built-up area)	
23	Signposting	Present Absent	
24	Direction	One direction (forth) One direction (back) Two directions	Yes*
25	Shoulder surface	No shoulder Hard shoulder Soft shoulder	
26	Lane width (in cm)	Number	Yes*
27	Intersection category (Sustainable Safety)	AR-AR (30 km/h inside BUA, 60 km/h outside) AR-DR (30x50 km/h; 60x80 km/h) DR-DR (50x50 km/h; 80x80 km/h; 50x80 km/h) DR-TR (50x70 km/h; 50 x motorway (MW); 80xMW) TR-TR (interchange; MWxMW)	

Table C.2 continued from previous page

	Characteristic	Input	Used
28	Priority	Right has priority Priority arranged Traffic signal control (traffic lights) Roundabout old style (traffic on arm has priority) Exit Other	
29	Arms	Number	
30	Crossing	None Pedestrian crossing/zebra (opt. with bicycle/moped) Channel markings (bicycle/moped and/or pedestrian)	

* For data filtering and preparing

D
Infrastructure data comparison

Several statistics regarding traffic safety in the municipalities in the Netherlands have also been gathered in the Road Safety Comparison Tool (Verkeersveiligheidsvergelijker) by Fietsersbond, SWOV and VVN (2020). This tool shows data about traffic deaths, reports by inhabitants, cycling infrastructure, the road surface and lighting, policies, and background information. The infrastructure data are taken from the route planner by the Fietsersbond[1] in 2019, and concern roads with speed limits of 50 and 80 km/h. The details about streets are maintained by volunteers that mostly know the situations they assess. As such, it is deemed useful to take a look at the similarities (or differences) between the shares of this tool and the full data set available for this study, and see if it is comparable and thus more reliable. Keep in mind that only data about bicycle tracks are shown in the tool. Therefore, also only numbers from tracks in the data set are used.

Table D.1: Comparison of tracks in the data set with the Road Safety Comparison Tool (Fietsersbond, SWOV and VVN, 2020)

<table>
<tr><th>Category</th><th>Subcategory</th><th>Data set (%)</th><th>Subcategory</th><th>Tool (%)</th><th>Difference (p.p.)</th></tr>
<tr><td rowspan="2">Track type</td><td>Bicycle tracks</td><td>77.16</td><td>Bicycle tracks</td><td>70.10</td><td>-7.06</td></tr>
<tr><td>Bicycle-moped tracks</td><td>22.84</td><td>Bicycle-moped tracks</td><td>29.90</td><td>7.06</td></tr>
<tr><td rowspan="3">Track directions</td><td>One-way</td><td>67.44</td><td>One-way</td><td>39.10</td><td>-28.34</td></tr>
<tr><td>Two-way</td><td>32.38</td><td>Two-way</td><td>60.90</td><td>28.52</td></tr>
<tr><td>Unknown</td><td>0.18</td><td>Unknown</td><td>0.00</td><td>-0.18</td></tr>
<tr><td rowspan="2">Access restriction on carriageway</td><td>Yes</td><td>1.58</td><td>Yes</td><td>80.80</td><td>79.22</td></tr>
<tr><td>No</td><td>98.42</td><td>No</td><td>19.20</td><td>-79.22</td></tr>
<tr><td rowspan="5">Track pavement type</td><td>Closed</td><td>71.24</td><td>Asphalt/concrete</td><td>71.00</td><td>-0.24</td></tr>
<tr><td rowspan="3">Open</td><td rowspan="3">28.34</td><td>Paving stones</td><td>18.30</td><td rowspan="3">0.66</td></tr>
<tr><td>Bricks</td><td>3.40</td></tr>
<tr><td>Other</td><td>7.30</td></tr>
<tr><td>Unknown</td><td>0.42</td><td>Unknown</td><td>0.00</td><td>-0.42</td></tr>
<tr><td rowspan="5">Track pavement quality</td><td rowspan="2">Sufficient</td><td rowspan="2">70.94</td><td>Good</td><td>64.10</td><td rowspan="2">21.36</td></tr>
<tr><td>Sufficient</td><td>28.20</td></tr>
<tr><td>Point of attention</td><td>25.51</td><td rowspan="2">Bad</td><td rowspan="2">0.90</td><td rowspan="2">-25.13</td></tr>
<tr><td>Bottleneck</td><td>0.51</td></tr>
<tr><td>Unknown</td><td>3.04</td><td>Unknown</td><td>6.80</td><td>3.76</td></tr>
<tr><td rowspan="4">Track street lighting</td><td rowspan="2">Yes</td><td rowspan="2">91.15</td><td>Good</td><td>81.30</td><td rowspan="2">-8.95</td></tr>
<tr><td>Limited</td><td>0.90</td></tr>
<tr><td>No</td><td>8.20</td><td>None</td><td>11.00</td><td>2.80</td></tr>
<tr><td>Unknown</td><td>0.65</td><td>Unknown</td><td>6.80</td><td>6.15</td></tr>
</table>

[1] https://routeplanner.fietsersbond.nl/

From the last column in Table D.1, that shows the difference between data in the tool and the used data set, the biggest changes are found in the number of directions on bicycle tracks, the presence of an access restriction for cyclists on carriageways and the pavement quality of bicycle tracks. The other categories are much closer and more or less the same. The number of directions is something that cannot be explained, however. For the access restriction – present in the regular road infrastructure assessment, but included here as well due to its presence from both sources – this could be because the assessment for the data set only noted whether a sign was present. For the tool, it might be that an access restriction automatically was in place when another facility for cyclists was present. For the analysis, only the best option to cycle is taken into account, so this does not matter. The track pavement quality difference at first seems big, but when taking into account that a point of attention actually has a low degree of cracks, holes and bumps and no acute danger of getting out of balance, it might as well be on the same side as Sufficient in the tool. Concluding, the infrastructure data set thus is fairly representative of the situation that will be studied (considering only bicycle tracks).

E
Correlations

Correlations between variables were used in the decision to include or exclude certain combinations of variables. In Table E.1, correlations for the composed obstacle variables of the filtered data set used for main effects are shown. The intensities were already put in natural logarithms. Dummies for the on-path obstacles medians and bollards were not included here, since these are very general. Variables for the type of shoulders and transitions were not created. With several absolute correlations being higher than 0.5, a value generally considered moderate to high (Hinkle et al., 2003), some combinations were excluded in the models. The quality of shoulders was only used when the obstacle-free zone share for the closest obstacles adjacent to the road surface was present. Obviously, only one obstacle-free zone share was included, ruling out high correlations between these variables. Intensities were both kept, since these are control variables.

The variable abbreviations mean the following:

- I: motor vehicle (mv) and cyclist (cyc) intensities
- S_{BN}: share of shoulder quality that is a bottleneck
- T_{BN}: share of transition quality that is a bottleneck
- OFZ: share of obstacle-free zones (adjacent and at 50 centimetres, 1 metre or 2 metres)
- med and bol: share of medians and bollards

Table E.1: Correlations between obstacle variables

	I_{mv}	I_{cyc}	S_{BN}	T_{BN}	OFZ_{adj}	OFZ_{050}	OFZ_{100}	OFZ_{200}	med	bol
I_{mv}	1	.667**	0.034	-.271**	-.284**	-0.089	0.051	.139*	0.097	0.062
I_{cyc}	.667**	1	.135*	-.268**	-.254**	-0.035	.154**	.151**	.127*	0.013
S_{BN}	0.034	0.135*	1	0.082	.300**	.623**	.976**	.603**	0.039	0.021
T_{BN}	-.271**	-.268**	0.082	1	.602**	.256**	0.058	0.008	0.010	0.040
OFZ_{adj}	-.284**	-.254**	.300**	.602**	1	.552**	.290**	.130*	-0.005	0.019
OFZ_{050}	-0.089	-0.035	.623**	.256**	.552**	1	.627**	.384**	0.016	0.070
OFZ_{100}	0.051	.154**	.976**	0.058	.290**	.627**	1	.620**	0.065	0.019
OFZ_{200}	.139*	.151**	.603**	0.008	.130*	.384**	.620**	1	0.092	-0.014
med	0.097	.127*	0.039	0.010	-0.005	0.016	0.065	0.092	1	-0.033
bol	0.062	0.013	0.021	0.040	0.019	0.070	0.019	-0.014	-0.033	1

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

The correlations between width variables of the filtered data set for main effects at a dominance of 75% of facility types are found in Table E.2. Since the values this time are fairly low, no combinations were excluded. Again, intensities were both kept due to those being control variables.

The variable abbreviations mean the following:

- I: motor vehicle (mv) and cyclist (cyc) intensities
- W: mean (M) and standard deviation (S) of the width of the dominating type in a street
- $DUMMY$: dummy for the three cyclist facility types (L for lane, $T1$ for one-way track and $T2$ for two-way track) next to the carriageway
- nar: narrowing share

Table E.2: Correlations between width variables

	I_{mv}	I_{cyc}	W_M	W_S	$DUMMY_L$	$DUMMY_{T1}$	$DUMMY_T2$	nar
I_{mv}	1	.687**	.454**	0.071	0.192**	.396**	0.107	-0.041
I_{cyc}	.687**	1	.406**	0.032	.266**	.372**	0.075	0.100
W_M	.454**	.406**	1	.230**	0.002	.511**	0.076	-0.038
W_S	0.071	0.032	.230**	1	-0.093	-0.081	-0.113	0.018
$DUMMY_L$	.192**	.266**	0.002	-0.093	1	-.305**	-.161**	0.023
$DUMMY_{T1}$	.396**	.372**	.511**	-0.081	-.305**	1	-.286**	-0.050
$DUMMY_{T2}$	0.107	0.075	0.076	-0.113	-.161**	-.286**	1	-0.067
nar	-0.041	0.100	-0.038	0.018	0.023	-0.050	-0.067	1

** Correlation is significant at the 0.01 level (2-tailed).

F

Determination width of carriageway

The flowcharts that show how the width of the carriageway was determined are depicted in Figure F.1 and F.2. The first only shows the steps for the left side, but the same can be constructed for the other side. For the sake of keeping the overview, Figure F.1 was split into two columns in the spreadsheet that was used for the analysis, leading to a value for either street sections where one side had its type reported as carriageway (orange diamonds), or both (yellow diamonds). The generic approach there was to subtract the width of an average car including small deviations from its path from the available carriageway width. On a 30 km/h road, this was found to be 2.40 metres (CROW, 2018b). While the average speeds in this study are 50 km/h, this width was chosen simply to arrive at a remaining cyclist width, where in the worst case cars have to slow down a bit to pass a cyclist. Some steps in the flowchart might seem strange, but since the data set did not always contain values for either of the used columns, the extra steps had to be incorporated. As the assessment of motor vehicle infrastructure contained valuable information for these steps, data from there was used as well (as indicated in Appendix C). The extra filters in the flowchart (blue and red diamonds), determining whether both sides are a carriageway or a continuous stretch, are furthermore found in Figure F.2.

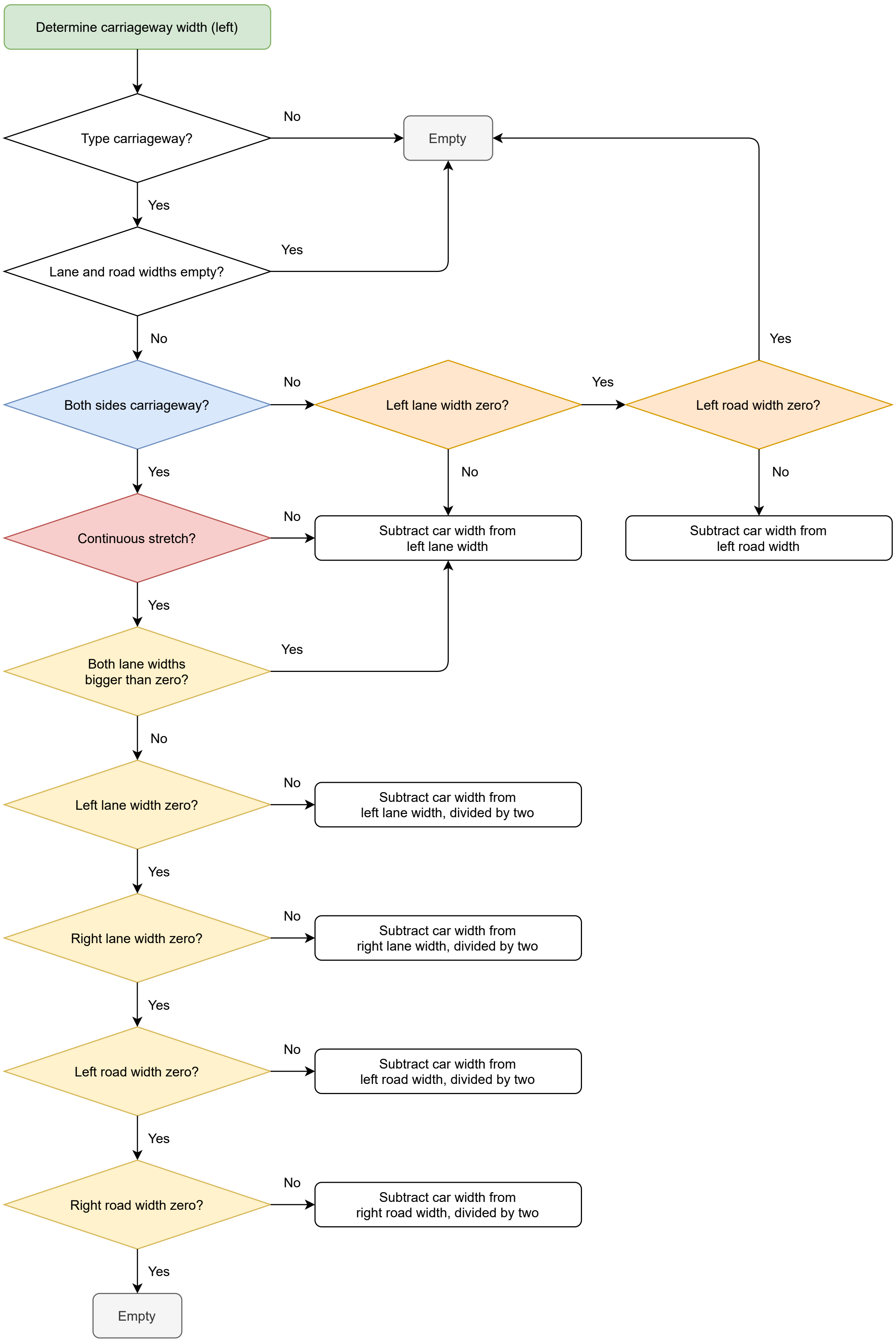


Figure F.1: Flowchart for determining carriageway width for cyclists (left side of the road)

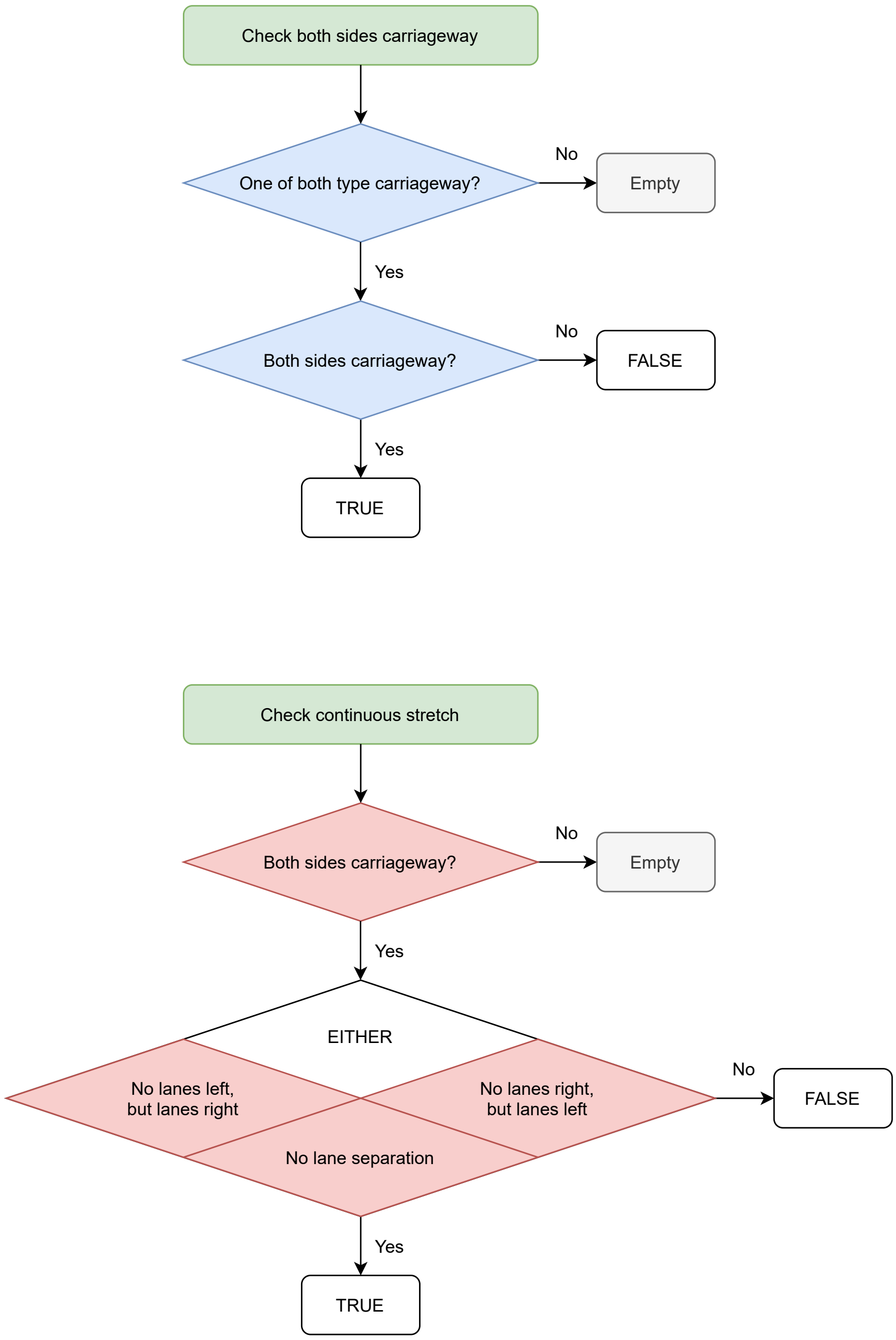


Figure F.2: Flowchart for carriageway checks: both sides are a carriageway (top) and the carriageways are a continuous stretch (bottom)

G

Sunset and sunrise

The table on the next page (in Dutch) shows the time of sunset and sunrise in the Netherlands for 2020, that were used to analyse a possible relation between light conditions with bicycle crashes (KNMI, 2019).

Koninklijk Nederlands
Meteorologisch Instituut
Ministerie van Infrastructuur en Waterstaat

tijden van zonopkomst en –ondergang 2020

KNMI

Bezoekadres
Utrechtseweg 297
3731 GA De Bilt
Postadres
Postbus 201
3730 AE De Bilt
T 030-220 69 11
F 030-221 04 07
www.knmi.nl
klimaatdesk@knmi.nl

Datum
24 september 2019

52°00' noorderbreedte en 5°00 oosterlengte

	januari		februari		maart		april		mei		juni		juli		augustus		september		oktober		november		december		
dag	*op*	*onder*	*op*	*onder*	*op*	*onder*	*op*	*onder*	*op*	*onder*	*op*	*onder*	*op*	*onder*	*op*	*onder*	*op*	*onder*	*op*	*onder*	*op*	*onder*	*op*	*onder*	*dag*
01	08.48	16.39	08.21	17.27	07.25	18.20	07.14	20.14	06.10	21.06	05.26	21.51	05.25	22.03	06.02	21.29	06.52	20.26	07.41	19.17	07.36	17.11	08.26	16.32	01
02	08.48	16.40	08.19	17.29	07.23	18.22	07.12	20.16	06.08	21.07	05.25	21.52	05.25	22.03	06.04	21.28	06.54	20.24	07.43	19.14	07.37	17.09	08.27	16.31	02
03	08.48	16.41	08.18	17.31	07.21	18.24	07.10	20.18	06.06	21.09	05.24	21.53	05.26	22.02	06.05	21.26	06.56	20.22	07.45	19.12	07.39	17.07	08.29	16.31	03
04	08.48	16.42	08.16	17.32	07.18	18.26	07.07	20.20	06.04	21.11	05.24	21.54	05.27	22.02	06.07	21.24	06.57	20.19	07.47	19.10	07.41	17.06	08.30	16.30	04
05	08.47	16.43	08.14	17.34	07.16	18.28	07.05	20.21	06.02	21.12	05.23	21.55	05.28	22.01	06.09	21.22	06.59	20.17	07.48	19.07	07.43	17.04	08.31	16.30	05
06	08.47	16.44	08.13	17.36	07.14	18.29	07.03	20.23	06.00	21.14	05.22	21.56	05.29	22.00	06.10	21.20	07.00	20.15	07.50	19.05	07.45	17.02	08.33	16.29	06
07	08.47	16.46	08.11	17.38	07.12	18.31	07.00	20.25	05.59	21.15	05.22	21.57	05.30	22.00	06.12	21.19	07.02	20.13	07.52	19.03	07.46	17.00	08.34	16.29	07
08	08.46	16.47	08.09	17.40	07.09	18.33	06.58	20.26	05.57	21.17	05.21	21.57	05.31	21.59	06.13	21.17	07.04	20.10	07.53	19.01	07.48	16.59	08.35	16.29	08
09	08.46	16.48	08.07	17.42	07.07	18.35	06.56	20.28	05.55	21.19	05.21	21.58	05.32	21.58	06.15	21.15	07.05	20.08	07.55	18.58	07.50	16.57	08.36	16.29	09
10	08.45	16.50	08.06	17.44	07.05	18.36	06.54	20.30	05.54	21.20	05.21	21.59	05.33	21.58	06.17	21.13	07.07	20.06	07.57	18.56	07.52	16.56	08.37	16.28	10
11	08.45	16.51	08.04	17.45	07.03	18.38	06.51	20.32	05.52	21.22	05.20	22.00	05.34	21.57	06.18	21.11	07.09	20.03	07.58	18.54	07.53	16.54	08.38	16.28	11
12	08.44	16.53	08.02	17.47	07.00	18.40	06.49	20.33	05.50	21.23	05.20	22.00	05.35	21.56	06.20	21.09	07.10	20.01	08.00	18.52	07.55	16.53	08.39	16.28	12
13	08.43	16.54	08.00	17.49	06.58	18.42	06.47	20.35	05.49	21.25	05.20	22.01	05.36	21.55	06.21	21.07	07.12	19.59	08.02	18.50	07.57	16.51	08.40	16.28	13
14	08.42	16.56	07.58	17.51	06.56	18.43	06.45	20.37	05.47	21.27	05.20	22.01	05.37	21.54	06.23	21.05	07.13	19.56	08.04	18.47	07.59	16.50	08.41	16.28	14
15	08.42	16.57	07.56	17.53	06.53	18.45	06.43	20.38	05.46	21.28	05.19	22.02	05.38	21.53	06.25	21.03	07.15	19.54	08.05	18.45	08.00	16.48	08.42	16.29	15
16	08.41	16.59	07.54	17.55	06.51	18.47	06.40	20.40	05.44	21.30	05.19	22.02	05.40	21.52	06.26	21.01	07.17	19.52	08.07	18.43	08.02	16.47	08.43	16.29	16
17	08.40	17.00	07.52	17.57	06.49	18.49	06.38	20.42	05.43	21.31	05.19	22.03	05.41	21.51	06.28	20.59	07.18	19.49	08.09	18.41	08.04	16.46	08.44	16.29	17
18	08.39	17.02	07.50	17.58	06.47	18.50	06.36	20.44	05.41	21.33	05.19	22.03	05.42	21.50	06.30	20.57	07.20	19.47	08.11	18.39	08.06	16.44	08.44	16.29	18
19	08.38	17.04	07.48	18.00	06.44	18.52	06.34	20.45	05.40	21.34	05.20	22.03	05.44	21.48	06.31	20.55	07.22	19.45	08.12	18.37	08.07	16.43	08.45	16.30	19
20	08.37	17.05	07.46	18.02	06.42	18.54	06.32	20.47	05.39	21.36	05.20	22.04	05.45	21.47	06.33	20.53	07.23	19.42	08.14	18.35	08.09	16.42	08.45	16.30	20
21	08.36	17.07	07.44	18.04	06.40	18.56	06.30	20.49	05.37	21.37	05.20	22.04	05.46	21.46	06.34	20.51	07.25	19.40	08.16	18.32	08.11	16.41	08.46	16.31	21
22	08.35	17.09	07.42	18.06	06.37	18.57	06.28	20.50	05.36	21.38	05.20	22.04	05.48	21.44	06.36	20.48	07.27	19.38	08.18	18.30	08.12	16.40	08.46	16.31	22
23	08.33	17.11	07.40	18.08	06.35	18.59	06.26	20.52	05.35	21.40	05.20	22.04	05.49	21.43	06.38	20.46	07.28	19.35	08.19	18.28	08.14	16.39	08.47	16.32	23
24	08.32	17.12	07.38	18.10	06.33	19.01	06.24	20.54	05.34	21.41	05.21	22.04	05.51	21.42	06.39	20.44	07.30	19.33	08.21	18.26	08.16	16.38	08.47	16.32	24
25	08.31	17.14	07.36	18.11	06.30	19.02	06.21	20.55	05.32	21.42	05.21	22.04	05.52	21.40	06.41	20.42	07.32	19.31	**07.23**	**17.24**	08.17	16.37	08.48	16.33	**25**
26	08.30	17.16	07.34	18.13	06.28	19.04	06.19	20.57	05.31	21.44	05.22	22.04	05.53	21.39	06.43	20.40	07.33	19.28	07.25	17.22	08.19	16.36	08.48	16.34	26
27	08.28	17.18	07.32	18.15	06.26	19.06	06.17	20.59	05.30	21.45	05.22	22.04	05.55	21.37	06.44	20.37	07.35	19.26	07.27	17.20	08.20	16.35	08.48	16.35	27
28	08.27	17.19	07.29	18.17	06.23	19.08	06.15	21.01	05.29	21.46	05.23	22.04	05.56	21.36	06.46	20.35	07.37	19.24	07.28	17.18	08.22	16.34	08.48	16.36	28
29	08.25	17.21	07.27	18.19	**07.21**	**20.09**	06.14	21.02	05.28	21.47	05.23	22.04	05.58	21.34	06.47	20.33	07.38	19.21	07.30	17.16	08.23	16.33	08.48	16.36	29
30	08.24	17.23			07.19	20.11	06.12	21.04	05.27	21.49	05.24	22.03	05.59	21.33	06.49	20.31	07.40	19.19	07.32	17.15	08.25	16.33	08.48	16.37	30
31	08.22	17.25			07.16	20.13			05.27	21.50			06.01	21.31	06.51	20.29			07.34	17.13			08.48	16.38	31

vanaf 1 januari tot en met 28 maart Midden-Europese Tijd | vanaf 29 maart tot en met 24 oktober Midden-Europese Zomertijd | vanaf 25 oktober tot en met 31 december Midden-Europese Tijd

begin van de lente: 20 maart 04.50 MET
begin van de zomer: 20 juni 23.44 MEZT
begin van de herfst: 22 september 15.31 MEZT
begin van de winter: 21 december 11.02 MET

H

Intensity forms

The following formulas show the transformations made to the original Poisson/negative binomial regression models to arrive at a model that explains the crash densities in streets with the motor vehicle and cyclist intensities in these streets, to back-up the thought of using natural logarithms for the intensities to mimic the "safety in numbers" principle. Extra (here: irrelevant) variables are depicted as $b_k X_k$, and Z is the product of the intercept with the last variable in the last lines. As can be seen in Formulas H.8 and H.14, the influence of intensities on the density is inherently different due to putting them in logarithms (the first set of formulas) or not (the last set). Figures H.1 and H.2 show plots of them. Lines in the first picture from top to bottom are b's with values of 0.125, 0.25, 0.5, 1 and 1.5. Colours are the same in other plots. The weakened effect of influence with increasing values is clearly visible on the left if regression coefficients stay below 1, which has been the case every time models were run.

$$\ln(C) = b_0 + \ln(L) + b_1 \ln(I_{mv}) + b_2 \ln(I_{cyc}) + b_k X_k \tag{H.1}$$

$$\exp(\ln(C)) = \exp(b_0 + \ln(L) + b_1 \ln(I_{mv}) + b_2 \ln(I_{cyc}) + b_k X_k) \tag{H.2}$$

$$C = \exp(b_0) * \exp(\ln(L)) * \exp(b_1 \ln(I_{mv})) * \exp(b_2 \ln(I_{cyc})) * \exp(b_k X_k) \tag{H.3}$$

$$= \exp(b_0) * L * \exp(b_1 \ln(I_{mv})) * \exp(b_2 \ln(I_{cyc})) * \exp(b_k X_k) \tag{H.4}$$

$$C/L = \exp(b_0) * \exp(b_1 \ln(I_{mv})) * \exp(b_2 \ln(I_{cyc})) * \exp(b_k X_k) \tag{H.5}$$

$$= \exp(b_0) * \exp(\ln(I_{mv}^{b_1})) * \exp(\ln(I_{cyc}^{b_2})) * \exp(b_k X_k) \tag{H.6}$$

$$= \exp(b_0) * I_{mv}^{b_1} * I_{cyc}^{b_2} * \exp(b_k X_k) \tag{H.7}$$

$$= I_{mv}^{b_1} * I_{cyc}^{b_2} * Z \tag{H.8}$$

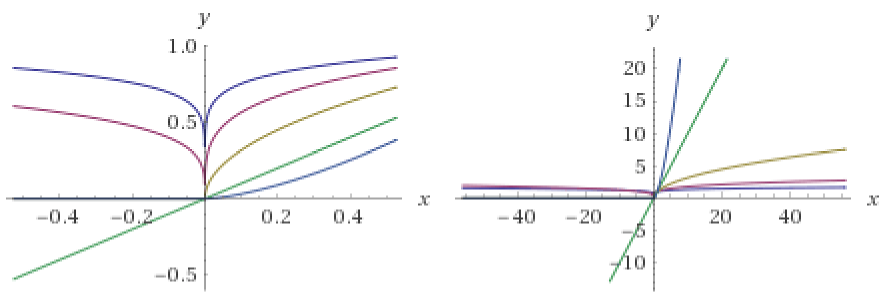


Figure H.1: Intensities powered with logarithm

$$\ln(C) = b_0 + \ln(L) + b_1 I_{mv} + b_2 I_{cyc} + b_k X_k \tag{H.9}$$

$$\exp(\ln(C)) = \exp(b_0 + \ln(L) + b_1 I_{mv} + b_2 I_{cyc} + b_k X_k) \tag{H.10}$$

$$C = \exp(b_0) * \exp(\ln(L)) * \exp(b_1 I_{mv}) * \exp(b_2 I_{cyc}) * \exp(b_k X_k) \tag{H.11}$$

$$= \exp(b_0) * L * \exp(b_1 I_{mv}) * \exp(b_2 I_{cyc}) * \exp(b_k X_k) \tag{H.12}$$

$$C/L = \exp(b_0) * \exp(b_1 I_{mv}) * \exp(b_2 I_{cyc}) * \exp(b_k X_k) \tag{H.13}$$

$$= \exp(b_1 I_{mv}) * \exp(b_2 I_{cyc}) * Z \tag{H.14}$$

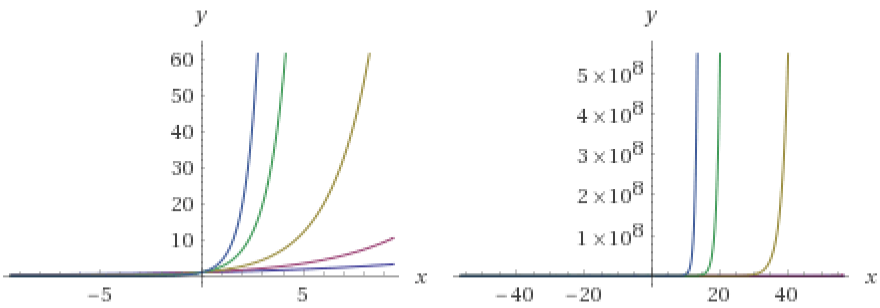


Figure H.2: Intensities exponentiated without logarithm

I

Additional interaction plots

The following figures contain additional interaction plots from models in the main report. Due to mean centring, all variables not involved in the interactions were kept at zero. The first two figures, I.1 and I.2 are plots of the interaction effects for visibility problems for medians and bollards (taken from Model O6). None of the parameter values is significant, although the one for medians inclines towards significance with a p-value of 0.06. The means are 0.0064 and 0.004, with a quarter of a standard deviation to the left (around 0%) for both and one standard deviation to the right (either 3.3% or 1.7%). Although neither effects are significant, both plots basically show what is expected. The higher the visibility problem share, the more bicycle crashes on a street occur. The lines not being (near-)parallel could be due to the exclusion of the main effect for the visibility problem share since it is a conditional variable, due to errors in measurement and even due to the low occurrence. Not much should be derived from these plots anyway.

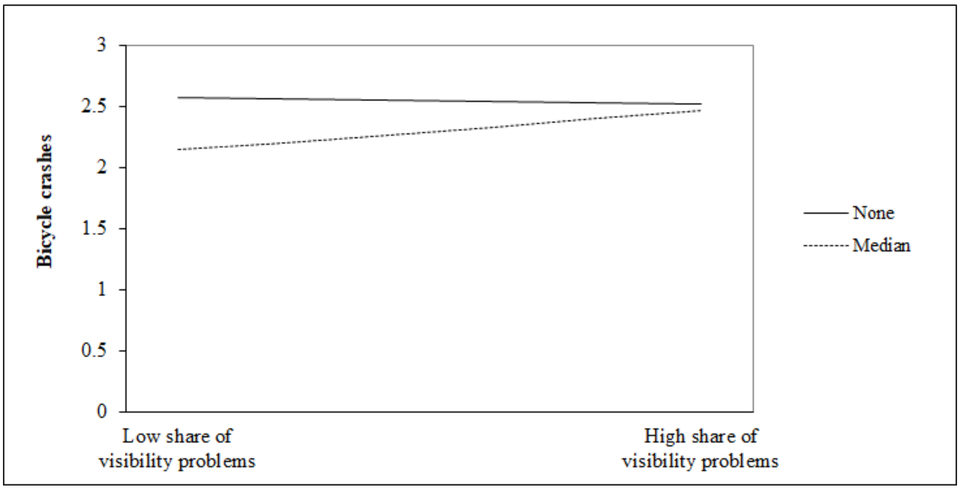


Figure I.1: Interaction plot for visibility of medians

The interaction plots for the width at different types of model CF are shown in Figures I.3, I.4 and I.5. Low and high widths are a standard deviation below and above the mean, resulting in 95 centimetres, 1.58 metre and 2.21 metres. They show the same results as for the width-crash models previously. An increase in width on lanes, one-way tracks and two-way tracks reduce bicycle crashes. An increase of the width on smaller sections with mainly lanes and one-way tracks has more impact than on wider sections. For two-way tracks, the lines are more or less parallel, indeed indicating the insignificance of the interaction effect.

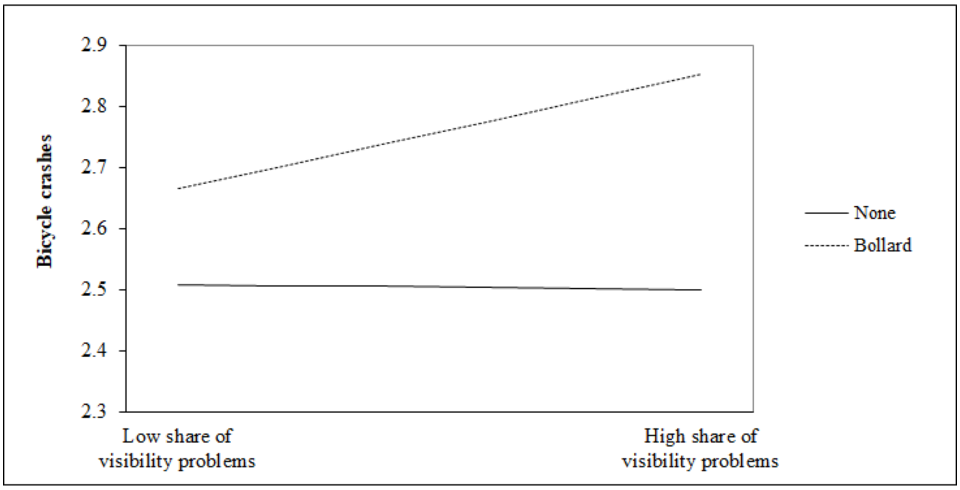


Figure I.2: Interaction plot for visibility of bollards

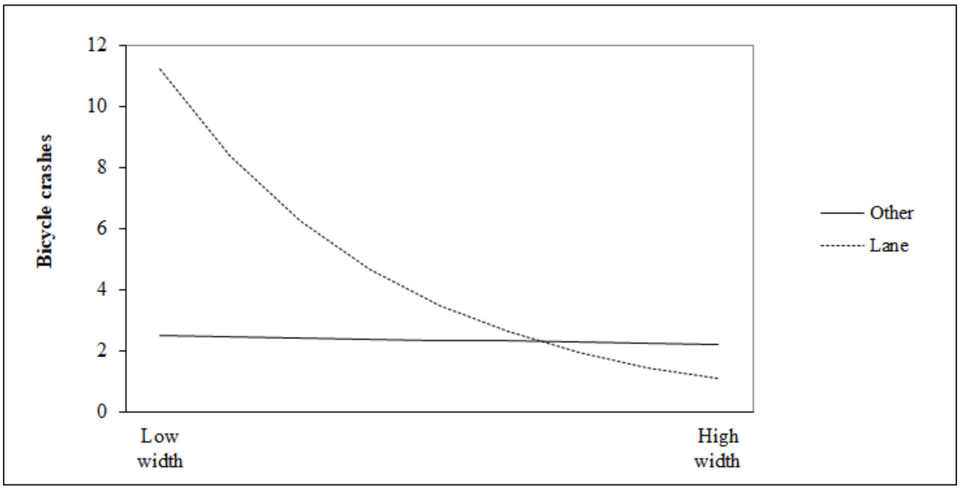


Figure I.3: Interaction plot for width and lane dummy in the combined model

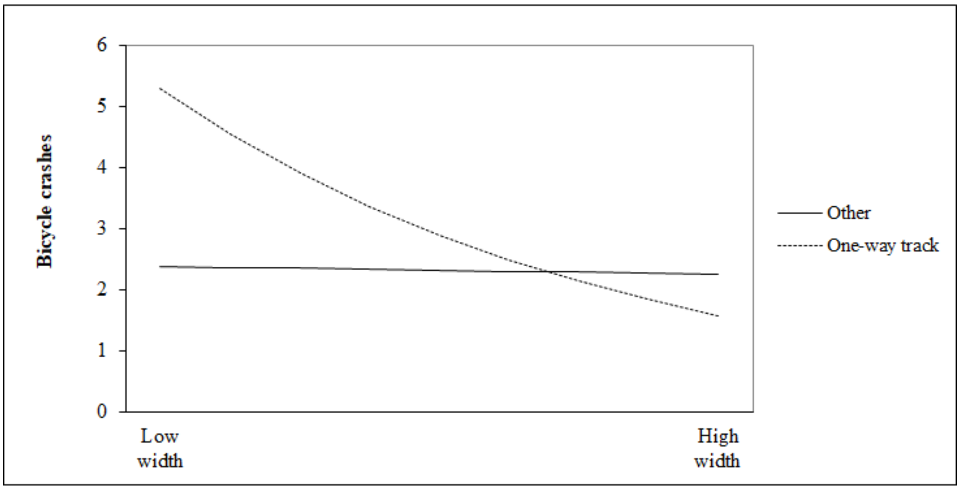


Figure I.4: Interaction plot for width and one-way track dummy in the combined model

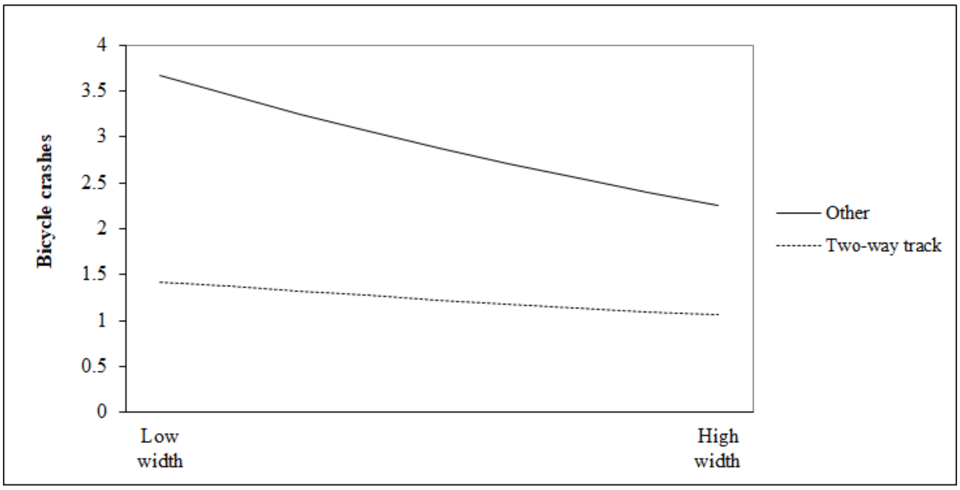


Figure I.5: Interaction plot for width and two-way track dummy in the combined model

Lastly, Figure I.6 contains a plot for the insignificant interaction effect of the transition bottleneck share at different widths of Model C3. The main effect was not significant either. The lines here were thus again expected to go parallel, and not head towards crossing each other, and even at the exact opposite slope. Now, namely they indicate that higher shares of bottlenecks are positive for bicycle crashes, which is not in line with expectations. However, it does not make sense to interpret this plot due to the insignificant variables and it is only useful to illustrate more can be done in this regard.

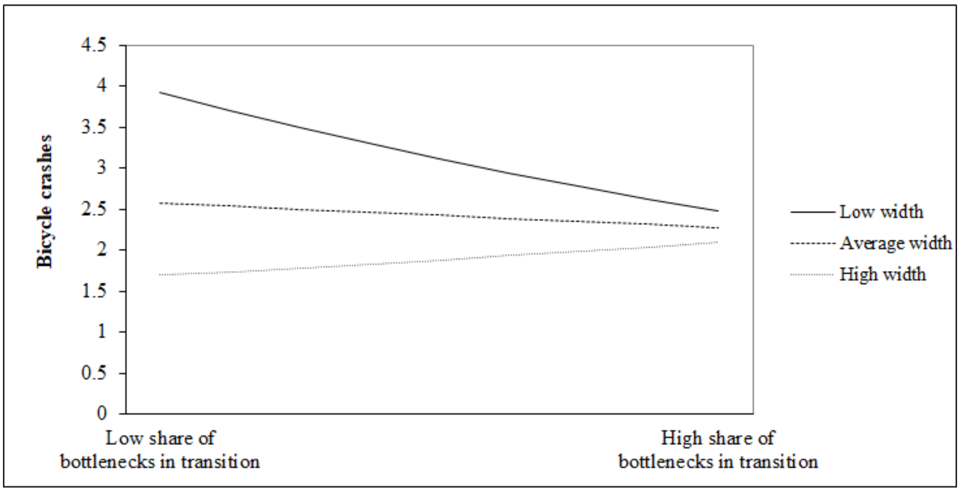


Figure I.6: Interaction plot for width and transition bottleneck share