Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 72 (2023) 367–374

Transportation Research Procedia

www.elsevier.com/locate/procedia

Transport Research Arena (TRA) Conference

Understanding the influence of accident risk and perceived safety on bicycle route choice

Stefan Huber[a]*, Paul Lindemann[a], Bettina Schröter[b]

[a]*Chair of Transport Ecology, Institute of Transport Planning and Road Traffic, Faculty of Transport and Traffic Science "Friedrich List", Technische Universität Dresden, 01062 Dresden, Germany*

[b]*Chair of Integrated Transport Planning und Traffic Engineering, Institute of Transport Planning and Road Traffic, Faculty of Transport and Traffic Science "Friedrich List", Technische Universität Dresden, 01062 Dresden, Germany*

Abstract

Little is known about the influence of accident risk and perceived safety on bicycle route choice. Do accident hot spots or points of 'perceived insecure' influence actual route choice of cyclists? To answer that question a GPS data set (~4,000 trips from ~170 study participants) and further data (infrastructure, operation, exposure, accidents reports, and survey data recording critical incidents) is used. The results reveal that a high accident risk along the route shows a significant negative but slight influence on route choice. In contrast, perceived safety does not significantly influence route choice. Other factors (e.g. existence or narrow width of cycling infrastructure, higher traffic volumes of bicycle traffic, signal controlled intersections) show significant influence and may represent perceived safety better than the used survey data.

Keywords: accident risk; perceived safety; bicycle route choice; GPS route data; multinomial logit

1. Introduction

Safety is a major issue in bicycle traffic and decreasing the accident risk is an important goal towards vision zero. In order to reach that goal and plan cycling infrastructure sufficiently, information on accident risk and perceived safety and its influence on cycling behavior is crucial. Although safety issues as well as bicycle route choice have been investigated within some studies separately, little is known about the influence of accident risk and perceived safety on bicycle route choice. Do, for instance, accident hot spots or points of "perceived insecure" influence actual

* Corresponding author. Tel.: +49 351 463-43156; fax: +49 351 463-37718.
E-mail address: stefan.huber1@tu-dresden.de

10.1016/j.trpro.2023.11.416

route choice of cyclists? Information from other studies provide raw explanation and interpretation on why and how cyclist choose their routes depending on safety issues.

Literature reveals that various international studies have been carried out in last decades to analyze bicycle route choice. First studies using revealed preference (RP) data were conducted in the 1980s/90s (see e.g., Lott et al. 1978 or Antonakos 1994). The number of studies increased in recent years (see e.g., the studies of Krizek et al. 2007, Menghini et al. 2010, Hood et al. 2011, Broach et al. 2012, Kang & Fricker 2013, Casello & Usyukov 2014, Kathri et al. 2016, Ton et al. 2017 and Chen et al. 2018). The mentioned studies investigate cyclists' route choice by using GPS data, which reveal the observed behavior of cyclists.

Although the focus of the studies, study area and the sample characteristics differ, the examination of the study results shows similar tendencies in bicycle route choice. The overall tendencies can be summarized as follows: Increasing distance and slope negatively influence route choice, whereas the existence of cycling infrastructure along a route as well as a good surface quality increase route choice probability. High traffic volumes and high levels of speed of motorized traffic along the route decrease route choice probability. Similar applies to a high number of intersections along a route.

However, safety is only indirectly taken into account in the studies. In most cases, the authors argue that, for example, the speed or volume of motorized traffic influences objective safety and the perception of safety along a route and, thus, safety has a considerable influence on route choice. This assumption is hardly doubted but safety is not directly considered and analyzed in the studies. Therefore, its influence has not been quantified yet. Influencing factors such as annual average daily traffic (ADT), heavy traffic shares or speeds of motorized traffic merrily serve as a kind of proxy in the studies. Hence, the influence of safety is interpreted to some extend but not quantified. The studies of Kang and Fricker (2013), Casello and Usyukov (2014) or Broach et al. (2012) exemplify that.

The study by Kang and Fricker (2013) represents a good example of studies that attempt to analyze the influence of safety on route choice indirectly via non-traffic-related influencing factors. The authors examine route choice focusing primarily on the decision of cyclists between riding on the pavement/cycle path (off-street) or using the road (on-street). Their research is based on surveys conducted on the campus of Purdue University in Indiana (USA). Students were asked about their routes and cycle path use when they arrived on campus. Thus, 178 cyclists were surveyed in this way from 2006 to 2008 collecting 931 trips. Based on the survey and other secondary data, Kang and Fricker (2013) conducted a statistical estimation to analyze the influence of person-specific characteristics (e.g. age, gender, wearing a helmet), environment-specific variables (e.g. time of day, wind speed, temperature) and route-specific variables (e.g. road class, number of lanes, gradients, stationary traffic facilities, ADT, surface conditions) on route choice. The result of their analyses is that cyclists generally prefer riding on the road (on-street). On the other hand, on large roads or roads with high ADT, they tend to use separate cycle lanes (off-street), which are also generally preferred over cycle lanes. Kang and Fricker (2013) finally conclude that there is a correlation between the use of specific cycling infrastructure (off-street versus on-street) and the perception of safety and surface quality. Similar interpretations in relation to safety can be found in Casello and Usyukov (2014) or Broach et al. (2012).

The other studies mentioned above neither analyze safety issues in connection with route choice. There are currently no studies examining the influence of objective and perceived safety on route choice on the basis of RP data – i.e., in which explicit safety variables were included into analyses. Thus, there are hardly any representative findings on safety factors influencing cyclists' route choice. The studies show that identified influencing factors are mainly based on comfort criteria (e.g., distances, slopes, surface characteristics) or other factors (e.g. type of infrastructure, speed and volume of motorized traffic) that are related to safety. However, there has been no analysis on how accident risk and perceived safety directly influence bicycle route choice based on real observation.

2. Data and methods

2.1. Route data

Bicycle route data reveal the choice behavior of cyclist. For the presented study, route data from a further research project (RadVerS) carried out by the Technical University Dresden were used. Within the mentioned project GPS route data have been collected within an extensive field study covering the city area of Dresden (Germany). The data set was recorded by 187 volunteer cyclists, which participated in the RadVerS bicycle research project. Participants were selected in terms of age and gender distribution corresponding to the population of the city. The participants were selected out of 10,000 people taking part in an online survey to determine types of cyclists (Lißner et al. 2020). Participants are composed of 80 female and 100 male riders aging from 16 to 88 years. Time of data collection was June/July 2018. The cyclists used an adapted version of the *Cyface* smartphone application (for iOS and Android) to track their rides. Beside the GPS data (latitude, longitude, timestamp), further data such as speed and accuracy were recorded and used for data processing. The data set initially contained more than 5,000 uploaded GPS tracks. Conducting a comprehensive data processing (filtering, smoothing, segmentation and mode choice recognition) around 4,000 trips were finally used for model estimation. Further information on data collection and processing can be found in Lißner et al. (2020) and Lißner & Huber (2021).

2.2. Transport supply data

In order to determine route characteristics, which are needed for model estimation, transport supply data is essential. Literature review shows that there is a huge variety of factors influencing bicycle route choice. The following factors have been defined and used as independent variables within analysis (precise variables in brackets):

- Distance (length of a route [km])
- Slope (maximum slope on the route [%] + share of the route with a slope < 2% [%])
- Existence of cycling infrastructure along the route (share [%])
- Surface quality (share of the route with good surface quality, e.g., asphalt [%])
- Parking (share of route section with parking [%])
- Speed limit of motorized traffic (share of route sections with max. speed ≤30km/h [%])
- ADT of motorized traffic (mean ADT along the route [veh/d])
- Number of lanes for motorized traffic (share of route sections with max. 1 lane [%])
- Presence of other cyclists (ADT_{cycle} > 500 cyclists/d [%])
- Pedestrian density (share of route sections with low pedestrian density along the route [%])
- No of intersections distinguished by operation (intersections with and without traffic light [n/km])
- Accident risk (number of accidents along the route [n])
- Perceived safety (number of points of perceived insecure along the route [n])
- Land use (share of living area and green area along the route [%])

Some of the required data was already available and taken from existing sources (e.g., data from the open data platform of the City of Dresden). In some cases, however, no information or only incomplete data was available, so that the required data had to be collected. The starting point for the collection was an in-house communication from the city of Dresden (basic information, e .g. on cycling infrastructure), which was updated and supplemented by using satellite images (Google Earth), Google Street View and, in some cases, site inspections. Information on the variables listed above were matched to a transport supply network model covering the city area. Thus, information for more than 350km of the city network was available for analysis, which represents around 25% of the overall road network (1,400km) within the city area (Dresden 2020). This represents the main road network and all important bicycle routes or rather corridors within the city area.

2.3. Accident data

To determine the influence of accident risk and perceived safety on route choice, various data sets were considered. On the one hand, data from the official accident statistics, which enable the spatial localization of accident sites and the identification of accident types. They are used to depict accident risk (objective safety). The data is freely accessible via the so-called *accident atlas* of the federal state statistics (DESTATIS 2020). Only accidents involving cyclist were considered.

Information from mobility reports of the RadVerS project is used to depict perceived safety. Beside route tracking, participants were furthermore asked to complete a mobility survey within the survey period. The mobility reports contained 407 records on reported incidents and their spatial location in the city area.

The spatial localization of the accident sites (objective safety) and the locations with reported incidents (perceived safety) is particularly relevant for the analyses. Thus, accident sites and locations with reported incidents were located in the geographic information system QGIS and assigned to the edges and nodes, which allows the assignment to routes and alternatives in a further step.

2.4. Choice set generation

In order to be able to analyze bicycle route choice, route alternatives must be generated. Alternatives can be generated using different methods (e.g., best-path method, deterministic or stochastic multi-path method). For the present work, we used an extended penalty tube approach according to Chen et al. (2007) and Bader et al. (2011). The used approach can be described as follows: 1. A buffer with diameter *d* is created around the already existing (chosen) route. All edges within the buffer are added to a penalty tubes list. 2. The cost (here: travel time) of all network edges from the penalty tubes list is scaled up with the penalty factor *pf*. 3. A shortest path routing (here: Dijkstra algorithm) is carried out between the start and end point of the chosen route. The routing algorithm takes into account the scaled cost (travel time) and searches for the 'least cost' path between the origin and the destination. 4. The found alternative is added to the set of possible alternative routes. 5. Step 1 is repeated for the found alternative, adding another penalty tube. The approach repeats step 1 to 5 until *k* alternatives have been generated.

The optimal parameters for the approach were determined using a test data set. We used $d=100m$, $pf=2$ and $k=2$. All routes in the choice set are spatially very different from the chosen route, so that 'real alternatives' were generated to avoid the IIA problem in model analysis.

2.5. Route choice model

In order to analyze route choice a multinomial logistic regression model (MNL) has been estimated. Logistic regression analysis is based on discrete choice theory (see McFadden 1976 and Ben-Akiva & Lerman 1985). Model formulation is as follows:

$$v_i = \beta_0+\beta_{i1}*x_{i1}+\beta_2*x_{i2}+\ldots+\beta_{in}*x_{in} \tag{1}$$

where v_i is the cost of each alternative, β_0 is the model constant, x_{ni} is the value of the independent variable n (see section 2.2) per alternative i, and β_n is the coefficient of the independent variable n. Using this alternative-specific utility v_i, an alternative-specific exponientiated utility u_i is calculated using equation (2).

$$u_i \;=\; e^{v_i} \tag{2}$$

with u_i as the exponentiated alternative-specific utility, e as Euler's number and v_i as the cost of each alternative. In a final step, the choice probability p_i for an alternative is calculated using u_i of all alternatives within the choice set for each OD relation. Calculation follows equation (3).

$$p_i(u_{i-n}) \;=\; \frac{u_i}{\sum u_{i-n}} \tag{3}$$

In order to determine the model constant and the coefficients, the maximum likelihood estimation (MLE) was used. It estimates the constant and coefficients by determining the highest possible probabilities for the observed value (choice or no choice) and is a well-established method for model estimation in discrete choice analysis (Schlaich, J., 2010).

The examination of fundamental requirements (e.g., categorical scale level of dependent variable, independent variables are metric, large sample size, few or no outliers in the sample, no correlation of independent variables) revealed that the regression analysis can be carried out because the data comply with all premises.

The statistical analysis was carried out using Biogeme (BisonBiogeme version 2.6a from 2017), a program well established in traffic sciences for the analysis of discrete choice (see Bierlaire 2003 and Bierlaire 2021).

3. Results

Out of 187 participants who took part in the data collection, the route data of 167 participants could be used (loss due to data processing). More than half of the participants were men (n = 96). Gender distribution is illustrated in Fig. 1 (left). Around 56.8% of all trips were made by men and 43.2% by women. Accordingly, men and women used the bike with similar frequency. The average age of the cyclists in the sample is 44.3 years. The youngest participants were 16 years old. The oldest one was 88 years old. This results in a variance of 72 years. The standard deviation is 14.5 years. The age distribution of the sample is illustrated in Figure 1 (right).

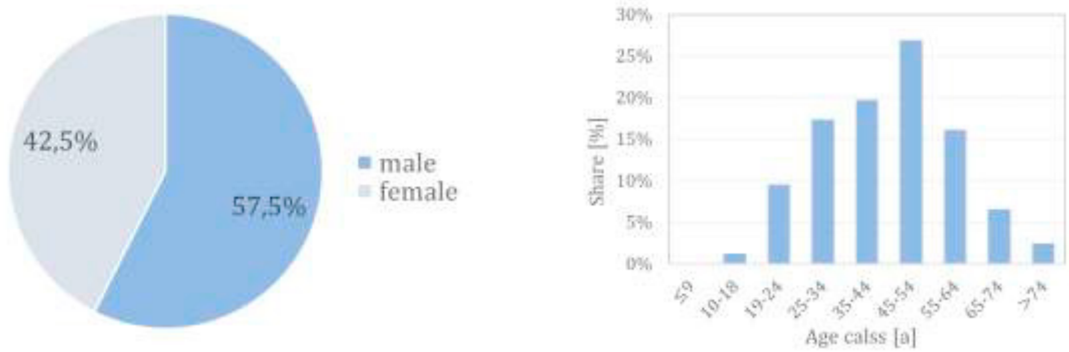

Fig. 1. Gender distribution (left) and age distribution (right) of study participants.

A total of 17,663 km was covered by the nearly 4,000 trips uploaded by the participants. This represents a considerable network coverage, as the total length of the cycle path network in Dresden is approx. 420 km (Dresden, 2020). Cyclist used the entire inner-city transport network. Highly frequented road sections mostly belong to the main transport network and the major cycling routes in the city. Road sections of the secondary network were less frequented.

The route choice behavior (see Table 1) can be summarized as follows: Although cyclists tend to stick to the shortest possible route, they are willing to take extra distance because of other crucial factors (e.g. cycling infrastructure or good pavement) that are much more important to them. Thus, the existence of cycling infrastructure, good pavement conditions (e.g., asphalt) as well as the existence of other cyclists along the route strongly and positively influence route choice probability. Less lanes for motorized traffic and an increasing proportion of route sections with low slopes (<2%) and green areas (e.g., parks) increase route choice probability, too. On the other hand, a negative influence is identified for increasing maximum slopes on the routes as well as for low pedestrian volumes. Other factors do influence route choice slightly or statistically not significantly (e.g., presence of parking or max. speed of motorized traffic ≤ 30km/h).

Most interesting is that both the accident risk and the perceived safety (number of reported incidents) show a marginal influence on route choice. The frequency of accidents along a route has a slightly negative effect on route choice probability (correlation coefficient = -0.004). The analyzed influence of the variable is statistically significant (p value = 0.000) so that the independent variable has a significant negative, albeit very weak, influence on route choice. The frequency of reported incidents (perceived safety) has a weak positive influence. If the number of reported incidents along the route increases, the probability also increases slightly (correlation coefficient = 0.037). The found correlation is also statistically significant (p-value = 0.000). The results of the route choice analysis are summarized in table 1.

At first glance, the results of the analysis with focus on safety are surprising. However, the influence of accident risk was found to be marginal. With regard to perceived safety, the results reveal a slight but statistically significant positive influence on route choice, which is unexpected so far (we assumed that cyclists avoid places where they do not feel safe while cycling). A possible explanation could be that reported incidents increasingly occur where people are cycling (high traffic volumes) so that routes inevitable pass these sights (e.g., if cyclists align their routes to the primary road network with high bicycle traffic volumes). Nevertheless, the influence of both accident risk and perceived safety turns out to be marginal, only. The route choice model shows a very good overall accuracy, as the pseudo R^2 statistics reveals high values (McFadden-R^2 = 0.494, Cox&Snell-R^2 = 0.662, Nagelkerke's R^2 = 0.745). Furthermore, the model shows a correct prediction of route choices (confusion matrix) of ~83%.

Table 1. Results of the multinomial logistic regression

Variable	Value	StdErr	t-test	p-value
Constant (shortest path)	0.789	0.050	15.710	0.000
Distance	0.255	0.041	6.240	0.000
Max. slope	-0.076	0.008	-10.120	0.000
Ratio of route length with slope < 2%	1.090	0.397	2.740	0.010
Cycling infrastructure	5.720	0.289	19.770	0.000
Asphalt pavement	2.860	0.282	10.120	0.000
Paring areas	0.382	0.281	1.360	0.170*
Max. speed of motorized traffic ≤ 30km/h	0.131	0.215	0.610	0.540*
ADT of motorized traffic	-0.000	0.000	-7.040	0.000
One lane for motorized traffic	1.880	0.294	6.420	0.000
ADT of bicycle traffic	2.410	0.206	11.700	0.000
Low pedestrian volumes	-1.290	0.524	-2.450	0.010
Intersections (rvl)	0.138	0.032	4.270	0.000
Traffic lights	0.270	0.075	3.580	0.000
Accident risk	-0.004	0.001	-5.220	0.000
Perceived safety	0.037	0.006	6.330	0.000
Land use 'residential area'	0.492	0.282	1.750	0.080*
Land use 'green area'	1.600	0.324	4.920	0.000

Figure 2 illustrates the explanatory power of the considered variables regarding bicycle route choice – or more precise: the ratio of increasing the models' quality (McFadden R^2) to its final level. The existence of bicycle infrastructure, high shares of road sections with low slopes, low maximum slopes on a route as well as good surface quality and a high bicycle ADT represent nearly 97% of route choice explanation. Thus, the analysis reveals that accident risk and perceived safety do not explain much of observed bicycle route choice decision. Cyclist obviously choose their route although there might be a safer or at least perceived safer alternative.

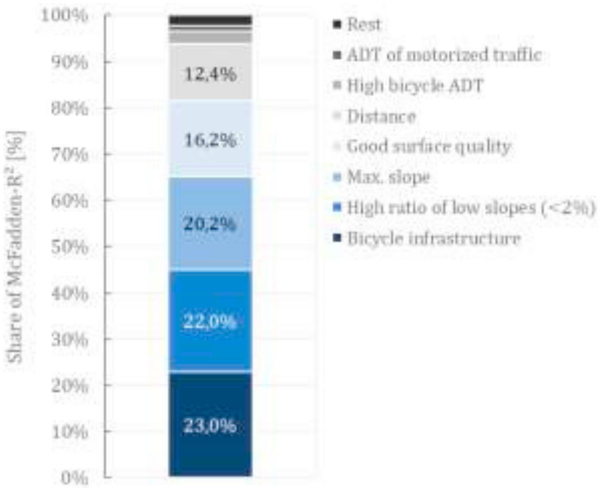

Fig. 2. Share of variables to model explanatory power

4. Discussion and conclusion

Both the influence of accidents risk (objective safety) and the number of reported incidents (perceived safety) do influence bicycle route choice of the cyclists in the sample marginally, only. The frequency of accidents along a route has a slightly negative effect on route choice so that the probability of choosing a route hardly decrease if the frequency of accidents increases along a route. Accidents are rare events and therefore cannot be remembered, let alone foreseen, by users at a particular location. However, this marginal effect is statistically significant, albeit quite weak. The influence of reported incidents on route choice was found slightly positive. If the number of points with reported incidents along a route increase, the probability of choosing this route increases slightly. This effect is significant, too.

The results of the analysis are surprising. Above all, the positive influence of points of reported incidents has not been expected in this manner. The influence of the accident frequency was assumed to be rather low or non-existent. The assumption was that only few cyclists take the accident locations within the urban area into account when choosing their route. In addition, statistically speaking, more accidents happen in places with higher traffic volumes (Alrutz et al. 2015), so that the result may be traced back to a positive correlation between route choice and accident frequency.

With regard to the influence of reported incidents, a negative correlation has been assumed at the beginning of the study. We assumed that cyclists avoid locations with more reported incidents compared to locations with few events. The positive influence of perceived safety (reported incidents) is very surprising. A possible explanation could be that reported incidents arose more frequently were people were cycling. Thus, there might be a spurious correlation between route choice and perceived safety. Since this is probably more likely to be the case in the main traffic network, these points are inevitably passed more frequently when driving through the city if cyclists use the main traffic network for orientation. These assumptions may explain the results of the analysis. In addition, the reported reported incidents come from the participants, who also recorded the routes. Hence, there is an inherent correlation between the data.

According to the overall results and the safety-specific findings, we can keep record that there seem to be more important factors than accident risk and perceived safety, which should be prioritized in planning and its´ sub disciplines (e.g., in modelling route choice of cyclist within traffic demand modelling). As the analysis was conducted for the case of Dresden, future work should conduct the analysis for other cities, which would enable to confirm/reject the results and discuss the influence of safety on observed route choice behavior in different contexts. The use of more data (route data as wells as accident data) could also help to extend the analysis and reclassify the results. A comparison of stated preferences and revealed preferences of route choice would be rewarding, too.

Acknowledgements

The authors greatly acknowledge the project funding by the Federal Ministry for Digital and Transport (BMDV) within the National Cycling Plan 2020 (NRVP).

References

Alrutz, Dankmar, et al. Einfluss von Radverkehrsaufkommen und Radverkehrsinfrastruktur auf das Unfallgeschehen. Forschungsbericht Nr. 29, Berlin: Unfallforschung der Versicherer. Gesamtverband der Deutschen Versicherungswirtschaft e. V. (GDV)., 2015.

Antonakos, C.L., 1994, Environmental and travel preferences of cyclists. Transportation Research Record 1438: 25-33

Bader, Roland, Jonathan Dees, Robert Geisberger, und Peter Sanders. „Alternative Route Graphs in Road Networks.“ International Conference on Theory and Practice of Algorithms in (Computer) Systems, 2011: 21-32.

Ben-Akiva, M., Lerman, S., 1985, Discrete choice analysis: theory and applications to travel demand. MIT Press Series in Transportation Studies, 1, pp. 59-99

Bierlaire, M., 2003, BIOGEME: A free package for the estimation of discrete choice models. Proceedings of the 3rd Swiss Transportation Research Conference

Bierlaire, M., 2021, BIOGEME Website. URL: https://biogeme.epfl.ch/

Broach, J., Dill, J. Gliebe, J., 2012. Where do cyclists ride? A route choice model developed with revealed preference GPS data, Transportation Research Part A: Policy and Practice, Volume 46, Issue 10, Pages 1730-1740, https://doi.org/10.1016/j.tra.2012.07.005

Casello, J., Usyukov, V., 2014, Modelling Cyclists Route Choice Based on GPS Data. Transportation Research Record, TRB Annual Meeting (2014)

Chen, P., Shen, Q., Childress, S., 2018, A GPS data-based analysis of built environment influences on bicyclist route preferences, International Journal of Sustainable Transportation, 12:3, 218-231, DOI: https://doi.org/10.1080/15568318.2017.1349222

Chen, Yanyan, Michael G. H. Bell, und Klaus Bogenberger. „Reliable Pretrip Multipath Planning and Dynamic Adaptation for a Centralized Road Navigation System." EEE Transactions on Intelligent Transportation Systems, vol. 8, no. 1, 2007: 14-20.

DESTATIS, 2020. Statistische Ämter des Bundes und der Länder – Unfallatlas. URL: https://unfallatlas.statistikportal.de

Dresden, 2020, State Capital Dresden, Landeshauptstadt Dresden - Verkehrsnetz und Verkehrsbelegung. Verkehrsnetze - Autobahnen, Straßen, Wege, Brücken und Tunnel." 19.. Mai 2020. URL: https://www.dresden.de/media/pdf/statistik/Statistik_3501_Verkehrsnetze_2014.pdf

Hood, J., Sall, E., Charlton, B., 2011. A GPS-based bicycle route choice model for San Francisco, California. Transp. Lett. 63–75. https://doi.org/10.3328/TL.2011.03.01. 63-75

Kang, L., Fricker, J.D., 2013, Bicyclist commuters' choice of on-street versus off-street route segments. Transportation, 40 (2013), pp. 887-902 https://doi.org/10.1007/s11116-013-9453-x

Kathri, R, Cherry, C.R., Nambisan, S.S., Han, L.D, 2016. Modeling route choice of utilitarian bikeshare users with GPS data. Transportation Research Record - Journal of the Transportation Research Board.

Krizek, K., El-Geneidy, A., Thompson, K., 2007, A detailed analysis of how an urban trail system affects cyclists' travel. Transportation, nº 34, pp. 611-624.

Lißner, S., Huber, S. 2021, Facing the needs for clean bicycle data – a bicycle-specific approach of GPS data processing. In: European Transport Research Review. DOI: https://doi.org/10.1186/s12544-020-00462-2

Lißner, S., Huber, S., Lindemann, P., Anke, J., Francke, A., 2020, GPS-data in bicycle planning: "Which cyclist leaves what kind of traces?" Results of a representative user study in Germany. In: Transportation Research Interdisciplinary Perspectives, Volume 7, http://dx.doi.org/10.1016/j.trip.2020.100192

Lott, D.Y., Tardiff, T., Lott, D.F., 1978, Evaluation by experienced riders of a new bicycle lane in an established bikeway system. Transportation Research Record 683: 40-46

McFadden, D., 1976, The mathematical theory of demand models. In: P. Stopher, A. Meyburg (Eds.), Behavioural Travel Demand Models, North Holland, Amsterdam, pp. 75-96

Menghini, G., Carrasco, N., Schüssler, N., Axhausen, K.W., 2010, Route choice of cyclists in Zurich. Transp. Res. A, 44 (9) (2010), pp. 754-776

Schlaich, J., 2010, Analyzing Route Choice Behavior with Mobile Phone Trajectories. Transportation Research Record: Journal of the Transportation Research Board. https://doi.org/10.3141/2157-10

Ton, D., Cats, O., Duives, D., Hoogendoorn, S., 2017. How do people cycle in Amsterdam? Estimating cyclists route choice determinants using GPS data from an urban area; Transportation Research Record Journal of the Transportation Research Board.