Abstract: The present study aims to deduce bikeability based on a collective understanding and provides a methodology to operationalize its calculation based on open data. The approach contains four steps building on each other and combines qualitative and quantitative methods. The first three steps include the definition and operationalization of the index. First, findings from the literature are condensed to determine relevant categories influencing bikeability. Second, an expert survey is conducted to estimate the importance of these categories to gain a common understanding of bikeability and merge the impacting factors. Third, the defined categories are calculated based on OpenStreetMap data and combined to a comprehensive spatial bikeability index in an automated workflow. The fourth step evaluates the proposed index using a multinomial logit mode choice model to derive the effects of bikeability on travel behavior. The expert process shows a stable interaction between the components defining bikeability, linking specific spatial characteristics of bikeability and associated components. Applied components are, in order of importance, biking facilities along main streets, street connectivity, the prevalence of neighborhood streets, green pathways and other cycle facilities, such as rental and repair facilities. The mode choice model shows a strong positive effect of a high bikeability along the route on choosing the bike as the preferred mode. This confirms that the bike friendliness on a route surrounding has a significant impact on the mode choice. Using universal open data and applying stable weighting in an automated workflow renders the approach of assessing urban bike-friendliness fully transferable and the results comparable. It, therefore, lays the foundation for various large-scale cross-sectional analyses.

Keywords: bikeability; cycling; active transport; built environment; infrastructure

1. Introduction

Cycling as an active mode of transport has advantages at societal level in respect of CO₂, air pollutant and noise emission as well as space requirements [1,2], and at an individual level when considering physical inactivity as a risk factor for common diseases of affluence [3–5]. Most recently, changing conditions in the transport sector caused by the COVID-19 pandemic led to a massive decline in public transport usage [6]. As a consequence, a healthy and environmentally friendly system of individual mobility with low access barriers is even more important. Earlier research found that more than half of the population can be defined as interested but concerned regarding using the bike as a mode of transport [7,8]. The choice of these people to use the bike is affected by policies. One key factor for encouraging bike traffic is the implementation of a bike-friendly built environment, i.e., road infrastructure.

Kellstedt et al. (2020) define bikeability as “the extent to which the actual and perceived environment is conducive and safe for bicycling” [9]. The idea of assessing the bikeability of an area and interrelating such to mobility behavior follows earlier approaches aiming to explain mobility patterns based on structural characteristics [10,11]. From a methodological perspective, the assessment of bikeability has developed from integrating the bicycle as
a depending variable into earlier frameworks [12], over-using strongly aggregated data such as bike lanes per square mile or per inhabitant in respect of whole cities or nations to explain cycling levels [13–16], to complex concepts considering various structural characteristics [17–21]. To date, many studies discussing bikeability are available [9,17,20–23]. Thereby, the secondary geodata used are mostly specific data obtained from municipal or national sources [13,24] or based on specific software to gather data [25]. At the same time, other studies still include considerable individual effort by raters in the field who collect associated data within an area instead of using secondary data [26–28]. Even though some studies follow a detailed data driven approach in local or regional case studies with specific operationalizations [29,30], across most prior studies, the included data are thematically similar overall but may be defined and operationalized differently.

Hence, studies provide valuable insights, but it is often not possible to transfer methodologies to other regions and the comparison of the results has its limitations. In addition, due to differences in data availability, it is almost impossible to apply detailed methodologies on a large-scale when crossing national borders. The lack of compatibility between such approaches highlights the importance of using a comparable and transferable index for the bikeability of urban neighborhoods using data with comprehensive coverage in different regions of the world [24]. Recent studies partly overcome this shortage by using open data but do not use a comprehensive definition of bikeability at the same time [22,31] while others elaborate detailed definitions, which partly rely on municipal data [17,19,32]. Earlier studies identified further research gaps, such as using reliable data and different types of infrastructures [33], integrating other kinds of infrastructure, such as bicycle shops or repair facilities [34], and collecting data with systematic consideration of relevant factors [35]. Unlike a coordinated and consensual consideration of impacting factors, various earlier studies improved our understanding significantly but use bike lanes as single parameters [13–16,31]. Other studies base weighting on single experiments [17,21] or normatively determine an overall framework [24]. As different types of infrastructure may have varying implications in different groups of cyclists [30,36], it is crucial to develop a consensual and integrative understanding of bikeability [9] and show its influence using statistical models.

The present contribution aims to address the following shortcomings in the field of bikeability identified by authors of prior studies: There is lack of a consensual definition [9] and a structured framework of influencing parameters [35]. Further infrastructures such as rental or repair facilities should be integrated [34]. In addition, using open geodata is desired in order to allow transferability and applicability [24]. Therefore, we aim to develop a method that includes relevant factors of bikeability in a common understanding and to set up a workflow using open data to allow the calculation of bikeability for any urban location. In addition, we aim to assesses the informative value of the resulting index by exemplary applying a multinomial logit mode choice model to describe the enhancements of bikeability on the decision of using a bike as the main transport mode in the city of Berlin.

The manuscript is structured as follows. First, we review the literature. Second, we describe the development of the methodology, containing an expert survey and the analyses of geodata. Afterwards we evaluate the method by quantifying the influences of the developed index in a mode choice model. Then, we describe and discuss the results and draw conclusions.

2. Background: Determinants of Bikeability

It is crucial when deriving bikeability to identify relevant influencing factors. Methodologies vary widely and a growing number of studies are available. Findings are generally obtained from cross-sectional studies [24,28], interventional studies [37–39], stated preference studies [40–43], revealed preference studies [44] and summarizing meta-analysis [10,45–48]. In the present contribution we focus on determinants evaluating the urban infrastructure, which is to some extent adaptable and lays in the responsibility of city or transport planning. In the following paragraphs, findings from previous research are grouped into five
categories: prevalence of neighborhood streets, street connectivity, biking facilities along main streets, green pathways, other facilities and further determinants. These are described below. For the following methodological approach, the most important determinants are used.

2.1. Road Types: Prevalence of Neighborhood Streets

There are various ways to operationalize the predominance of motorized traffic on different roads. Previous studies, using the number of car lanes [49], traffic volumes [50–52], permitted speed [53], road categories [54] or a combination of several characteristics [55], deliver varying results. Many interventional or cross-sectional studies prove the importance of calmer streets [56–58] while others do not [59]. Some revealed preference studies verify the preference for local roads over arterials [44,60] whereas other studies find diverging results [61] or no interrelations at all [62]. When questioning cyclists’ route preferences in stated preference studies, giving priority to calmer streets over main streets appears obvious [54] even though the preference for dedicated cycle infrastructure or off-street paths might be stronger [63].

2.2. Street Connectivity

Street connectivity or intersection density is an established infrastructural factor associated with mobility behavior [10,11,64]. Many studies prove the significance of higher intersection density on active transport routes [56,64–68]. When operationalizing intersection density as a directness of route in stated preference studies, low directness is found to be a hindering factor for cycling to work [52]. Less confident cyclists on the other hand prefer a low number of junctions [40].

2.3. Biking Facilities along Main Streets

Both stated preference [40,63] and revealed preference studies [44,69,70] conclude that cycle infrastructures are preferable to cycling in mixed traffic and can compensate for the adverse effects of busy streets. Many cross-sectional [13,14,16,71,72] or longitudinal interventional studies [73–75] also find correlations between the amount of cycle infrastructure and cycling frequency at different aggregation levels. In contrast, other cross-sectional [64] and revealed preference studies [62] refute such interrelations. One reason for inconsistent results is said to be not distinguishing adequately between different types of infrastructure in known studies [33].

2.4. Green Pathways

Findings related to cycling on green pathways appear inconsistent across all methodological approaches. Stated preference studies range from preferring off-street paths above all other routes [63] to off-street paths being less attractive than paths along arterials [54]. Similarly, revealed preference studies find no [62], low [61] or high [44,60,70] significance for choosing off-street paths. Cross-sectional studies provide contradictory findings ranging from a slightly positive [76] to no [56] or negative correlation [66]. This might be due to operationalization issues or to interrelations with other structural factors.

2.5. Other Cycle Facilities Such as Rental and Repair Facilities

As the number of bike-sharing systems increases, it is becoming increasingly relevant to analyze their impact on bikeability and cycling mode shares. Previous studies concluded that bike-sharing has a positive influence on cycling levels but needs further measures to become effective [77–79]. Few results illustrate the positive impact of other kinds of infrastructure, such as availability of repair facilities, air pumps or bicycle shops [79,80]. They are also mentioned as promising areas of interest and research subjects for further studies [34].
2.6. Further Determinants

In addition to the categories listed above, several further determinants influence bicycle usage. For example, the quality of service for complementing modes of transport, such as walking or public transport, is stated to interact with cycling levels [71,81–83]. Moreover, the natural environment, especially slopes, strongly influences local bike-friendliness [23,24]. In addition, further road characteristic like surface quality [63,84,85] or intersection design [86] may be crucial. When aiming to explain cycling levels it may also be beneficial to include land use or built environment parameters as these influence traffic demand in general [45,87,88]. Furthermore, indicators which are caused by motorized transport such as air pollution [89,90], traffic noise [91,92] or safety considerations/numbers of accidents [93,94] can play a role. Finally, indicators mentioned above cause the specification of depending factors like perceived safety [95,96] or comfort [97,98]. These also strongly depend on individual attitudes towards cycling [7]. Even though perceived safety is a major underlaying motive, research does not prove areas with higher bikeability are objectively safer [99].

3. Methodology

The approach combines different methods and contains four steps which subsequently build on each other. The methodological approach is displayed in Figure 1. Thereby, the first three steps refer to developing the methodology. First, literature is reviewed to determine factors influencing bikeability as described above. This provides a list of determinants influencing bikeability. Second, an expert survey is conducted to rate the importance of deduced categories to arrive at a conceptual bikeability weighting of the determinants. Third, open geodata are used to operationalize parameters for each influencing category and calculate individual values. Single parameters are merged to form an overall bikeability index by using the weighting determined. Fourth, using Berlin as an application example, a model is set up to evaluate the method developed using the bikeability index in a multinomial logit mode choice model.

![Figure 1. Methodological approach.](image)

3.1. Expert Survey: Gather Weighting of Determinants

A weighting is needed to merge different environmental parameters according to their specific importance to bikeability. Consulting the authors of many studies and other cycling-experts promises a more stable weighting than referring to a single element experiment. Therefore, an expert survey is conducted to gain a common understanding of bikeability.

The survey uses a website containing one main interactive module as shown in Figure 2. Each of the five influencing categories (see section background: determinants of bikeability) is represented by a slider control. Moving the value of one category scales the other categories accordingly to add up to 100 percent. Participants are asked to apply the tool, specifying a general weighting based on their professional assessment of bikeability. Participants are also invited to add an impacting factor if they feel a category is missing. Providing this option is important in order to allow for stating other influencing factors than those predefined.
A total of 141 experts were personally invited to take part in the survey. The experts were contacted by email between 14 September and 19 September, 2016. Those invited were corresponding authors of the studies reviewed above (74) and published in the last four years, members of the research group of Germany’s National Cycling Plan 2020 (32), and other well-known experts (35). Experts were assigned a personal code relating to a separate database containing information on the reason for selection, country of residence, profession (practical or research) and gender. When answering the survey, experts were asked to “imagine an urban neighborhood which generally invites bike riding in everyday traffic” and “weigh the importance of different elements of urban infrastructure for local bikeability”. The goal was stated as calculating a bikeability index using spatial data.

3.2. Geodata Analyses: Calculation of Parameters

To ensure the transferability of the proposed method, data from OpenStreetMap (OSM) are utilized to calculate built environment parameters [100]. OSM represents an ideal basis for this methodology due to its standardization, comprehensive coverage, high level of detail and freedom from financial and legal constraints [101–103]. Using OSM data, a workflow is set up to derive spatial parameters which can be calculated for any scale and extent. Each category is operationalized to a conclusive parameter allowing for various ways of tagging different types of infrastructure in the data. Different vector data sets, dividing the investigation area into grids or administrative zones (cells), can be loaded for spatial localization. Different administrative zones can also be extracted directly from OSM. Spatial indices are then calculated for each parameter. As the resulting values include different units and scales, they are standardized to permit unification into one overall bikeability index. Afterwards, the standardized values are merged using the weighting from the expert survey. Computing of individual indices for each category is described below.

To estimate road types, the road’s importance indicated in the network for motorized traffic from motorway to traffic-calmed street is used. It is assumed that the road’s importance to motorized traffic contrasts with the convenience of using that road by bicycle. This interrelation is implied in literature and the expert survey. All road segments are added up for each cell in the relevant category. The percentages of the three smaller road types (traffic-calmed, residential and tertiary streets) are added up to create an index and estimate the prevalence of small streets.

To estimate intersection density, intersections negotiable by bike are counted and the density per square kilometer is calculated. Each node in the data is analyzed. If a node is assigned to more than one highway, it is defined as a crossing street line. As intersec-
tions often contain a number of crossing street lines, a cluster algorithm based on the DBSCAN \cite{104} is used. Every cluster identified corresponds to one intersection.

To calculate the coverage of cycle facilities along main streets, the two main street categories usable by bicycle (primary and secondary) are analyzed. When calculating the coverage of biking facilities, both the variety of facilities that exist and different ways of tagging them are taken into account. Here, cycle lanes at street level and separated cycle tracks are considered by analyzing properties of the main street segment. Cycle tracks are also tagged as separate routes adjacent to the road when separated by grass or hedges. To capture these adjacent routes, a 10-meter buffer is placed around each street segment to identify cycle tracks running parallel to that segment. In each cell, the percentage covered by three types of facilities is calculated for both the types of streets and the direction of traffic. When added, the result is the percentage coverage of main streets with cycle infrastructure.

To approximate the prevalence of green pathways, all public green areas accessible by bike are taken into account. In this case, differently tagged areas often overlap. For instance, small-wooded areas within parks consist of two overlay polygons which are considered using an overlay algorithm. The area of public green spaces in each cell is added up and the percentage share of greenery in each cell is calculated. The actual length of trails is not taken into account as there is a significant difference between the density of tracks running through green spaces and local tagging behavior in the OpenStreetMap. In some areas every beaten track is included in the data. In others only official main pathways are tagged. Counting the length of paths would lead to distortions.

The category rental and repair facilities aggregates bike-sharing stations and bicycle shops providing repair or rental services in each cell. The amount is added up and the density per square mile is calculated.

3.3. Mode Choice Model: Influence of the Bikeability

In order to evaluate the influence of the calculated index on choosing the bike for a single trip a mode choice model has been applied. Basis of this analysis is the municipal household travel survey data (SrV) of the year 2008, which provides information about 73,667 valid trips within Berlin, 10,234 of them were covered by bicycle \cite{105}. In our case valid means that all trips start and end within the city border of Berlin, that they do not start and end in the same cell and that they do not include null values or values that are not plausible. The survey includes information about each trip (length, origin and destination area, trip purpose, main transport mode, etc.) and the conducting person (age, employment status, availability of a car or public transport ticket, etc.), aggregated to the spatial level of the 195 statistical units of Berlin (see Figure 3). Therefore, it is not possible to obtain the exact start and end points of single trips. Due to this limitation, center points of the respective districts have been calculated and used as origin and destination locations (see Figure 3a) as an approximation. Afterwards, a shortest path routing on the bike network has been applied between each pair of center points (see Figure 3b) and the obtained routes were subsequently intersected with the districts and associated bikeability values (see Figure 3c). Finally, the bikeability values of the crossed districts have been summed up and averaged for each trip.

In order to evaluate the influence of the calculated bikeability index on the choice of using the bike as primary mode instead of motorized transport for a respective trip a multinominal logit (MNL) model has been applied.

The MNL model is based on the assumption, that an individuum that faces a decision between distinct alternatives chooses the alternative that comes with the highest utility. The random utility theory, under the assumption of additive linearity, states that the utility (\(U\)) of one alternative is the sum of their single components (\(X\)) and an independent and identically distributed (i.i.d.) error term (\(\epsilon\)). The utility of one alternative can then be formulated as:

\[
U = X\beta + \epsilon
\]
where the parameters $\beta$ are to be estimated and reflects the influence of the respective components $X$ \cite{106,107}.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Process of averaging bikeability values of each trip. The figure shows all statistical units of the district “Friedrichshain-Kreuzberg” as an example. In this illustration one of the 38,025 regarded trips is visualized. The process includes the calculation of area centroids (a), the routing between one pair of centroids (b) and the averaging of bikeability values based on the crossed units (c). In this example the bikeability values of the regarded statistical units range between 0.02 and 0.97, the average value for the observed trip is 0.53.}
\end{figure}

The dataset used for estimation includes all trips of the trip data described above, which are made by car, bike, public transport or foot as the corresponding main mode. From this, the dataset was cleaned by removing trips with implausible travel speeds (e.g., car trips faster than 100 km/h on average or walking trips over 7 km/h) and trips with distances less than 500 meters, which then reduces the dataset to a final of 48,825 observations. The dataset contains several properties of each conducted trip (the original route and mode), which were stated in the survey. As the model requires information on the alternatives which were not chosen, we calculated these respective route characteristics of the alternatives for each trip using the modes that were not chosen in the observed original trip (e.g., travel times and costs for public transport, car or bicycle).

The components of the utility function used for the estimation consists of trip attributes as well as person attributes. The former includes mode specific cost and travel time of the trip and the bikeability index for bike trips, the latter contains the age of the respondents and the availability of a public transport season ticket. The car availability of participants is considered in the availability of alternatives, i.e., a person without access to a car will not be given an alternative to drive. All attributes used are shown in Table 1, enhanced with a short description and the minimum, maximum and mean value. It is important to note that the statistics are based on the values of all alternatives, including those which were not chosen.
Table 1. Attribute table.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>Age of the person observed in years</td>
<td>6</td>
<td>95</td>
<td>44.74</td>
</tr>
<tr>
<td>PT_TICKET</td>
<td>Public transport season pass available</td>
<td>0</td>
<td>1</td>
<td>0.48</td>
</tr>
<tr>
<td>COST PT</td>
<td>Costs of the trip with public transport in €</td>
<td>0.2</td>
<td>12.15</td>
<td>1.73</td>
</tr>
<tr>
<td>COST CAR</td>
<td>Costs of the trip by car in €</td>
<td>0.005</td>
<td>6.53</td>
<td>0.66</td>
</tr>
<tr>
<td>BIKEABILITY</td>
<td>Bikeability index of the route</td>
<td>−0.87</td>
<td>0.97</td>
<td>0.18</td>
</tr>
<tr>
<td>TIME_BIKE</td>
<td>Travel time by bike in minutes</td>
<td>2.08</td>
<td>130.01</td>
<td>14.39</td>
</tr>
<tr>
<td>TIME_CAR</td>
<td>Travel time by car in minutes</td>
<td>1.13</td>
<td>111.11</td>
<td>19.75</td>
</tr>
<tr>
<td>TIME_TRAIN</td>
<td>Travel time by public transport in minutes</td>
<td>3.95</td>
<td>741.2</td>
<td>68.73</td>
</tr>
<tr>
<td>TIME_WALK</td>
<td>Travel time by foot in minutes</td>
<td>8.33</td>
<td>133.33</td>
<td>24.36</td>
</tr>
</tbody>
</table>

4. Results

4.1. Expert Survey

With 57 valid responses, the return rate for the survey is 40.4%: one third from females and two thirds from males. 33 (58%) respondents are from Germany, 13 (23%) from other European countries and 11 (19%) from America. The profession for 44 (77%) participants is researcher while 13 (23%) are working in practice. One third of all respondents published at least one related peer-reviewed article in the last four years.

Table 2 shows the results of the participants’ assessment of bikeability. Biking facilities along main streets are the most important component of bikeability but also show the highest variations in ratings. Further components are, in order of importance, street connectivity, prevalence of neighborhood streets and green pathways. Other cycle facilities are less important. Apart from assessing the described categories, one third of respondents (19) added and rated an individual category. Six respondents mentioned parking. Surface quality was also mentioned three times. Other categories were entered only once and were partly outside the scope of factors influencing infrastructure. When mentioned, both parking and surface quality attracted relatively high values (18.9 or 20.6 percent, respectively, on average). As most experts did not feel a category in addition to the five predefined important, the mean value for the additional category stays low. Providing this option was important for the process to allow for further determinants. Results show that experts confirm the predefined categories and use the option of adding new categories rarely.

Table 2. Expert survey: resulting interaction of components of bikeability.

<table>
<thead>
<tr>
<th></th>
<th>Prevalence of Neighborhood Streets</th>
<th>Street Connectivity</th>
<th>Biking Facilities along Main Streets</th>
<th>Green Pathways</th>
<th>Other Cycle Facilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.17</td>
<td>0.23</td>
<td>0.28</td>
<td>0.16</td>
<td>0.08</td>
</tr>
<tr>
<td>Median</td>
<td>0.14</td>
<td>0.24</td>
<td>0.27</td>
<td>0.16</td>
<td>0.07</td>
</tr>
<tr>
<td>Standard-deviation</td>
<td>0.10</td>
<td>0.90</td>
<td>0.11</td>
<td>0.09</td>
<td>0.06</td>
</tr>
<tr>
<td>Min.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Max.</td>
<td>0.45</td>
<td>0.45</td>
<td>0.52</td>
<td>0.36</td>
<td>0.25</td>
</tr>
</tbody>
</table>

* percent missing up to 100: additional category.

Variances in the experts’ assessments were compared across various groups (gender, profession, location of residence), resulting in no substantial differences regarding the subgroups under consideration. Statistical tests found no significant differences regarding means and variance in any of the components across the groups. Table 3 compares the experts’ assessment in Europe and America as different regions of the world. As seen, there are no structural differences. Experts based in Europe rate the prevalence of neighborhood streets slightly higher while experts based in America value biking facilities along main streets a bit higher.
Table 3. Experts’ assessment in different geographical regions.

<table>
<thead>
<tr>
<th>Region</th>
<th>Prevalence of Neighborhood Streets</th>
<th>Street Connectivity</th>
<th>Biking Facilities along Main Streets</th>
<th>Green Pathways</th>
<th>Other Cycle Facilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>America</td>
<td>0.14</td>
<td>0.23</td>
<td>0.31</td>
<td>0.16</td>
<td>0.09</td>
</tr>
<tr>
<td>(n = 11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>0.17</td>
<td>0.23</td>
<td>0.28</td>
<td>0.15</td>
<td>0.08</td>
</tr>
<tr>
<td>(n = 46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* percent missing up to 100: additional category.

4.2. Characteristics of Parameters

Results are calculated using the example of Berlin, Germany, as proof of concept and to analyze intra-city characteristics and evaluate the methodology.

The structure of the network differs greatly when considering road types. Motorways and trunk roads form a loose network and are only present in a few districts. However, the percentage share of motorways in one particular cell is more than 45%. Primary streets appear as a star-shaped network with relatively low shares of 5% on average and a rather low variation. Secondary and tertiary streets account for substantial shares in the network with 17 and 11 percent, respectively, on average. Residential streets dominate most districts with 62% on average but ranging from 0 to 100%. Traffic-calmed streets account for low shares in many districts and also show dispersed distribution with high shares locally in a few districts. As shown in Table 4, small streets account for large shares in most districts.

Table 4 shows the statistical characteristics of all input parameters for bikeability. All values range widely.

Table 4. Statistical key figures of individual parameters.

<table>
<thead>
<tr>
<th></th>
<th>Prevalence of Neighborhood Streets (Percentage Share of All Street km)</th>
<th>Street Connectivity (Intersections per km²)</th>
<th>Biking Facilities along Main Streets (Percentage Coverage)</th>
<th>Green Pathways (Percentage in Area)</th>
<th>Other Cycle Facilities (Rental and Repair Facilities per km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.76</td>
<td>37.6</td>
<td>0.65</td>
<td>0.18</td>
<td>0.91</td>
</tr>
<tr>
<td>Median</td>
<td>0.81</td>
<td>35.9</td>
<td>0.69</td>
<td>0.11</td>
<td>0.16</td>
</tr>
<tr>
<td>Standard-deviation</td>
<td>0.18</td>
<td>22.2</td>
<td>0.28</td>
<td>0.21</td>
<td>1.78</td>
</tr>
<tr>
<td>Min.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>Max.</td>
<td>1</td>
<td>108.9</td>
<td>1</td>
<td>0.97</td>
<td>10.32</td>
</tr>
</tbody>
</table>

Figure 4 shows the spatial characteristics of the inner-city distribution of single values and the overall index based on Berlin. Darker colors indicate higher values. With regard to the individual parameters, the share of small streets (top left) shows a distribution with generally higher values in outer districts. The map of intersection density (top right) shows a centered distribution with the highest values in inner districts. The denser districts in the city center are characterized by a highly connected street network compared to lower intersection density in outer districts. The coverage of main streets with cycle facilities (middle left) appears in a highly dispersed distribution. The share of greenery (middle right) is highest in outer districts with the large green areas in the west and southeast easily recognizable. The spatial density image of rental and repair facilities (bottom left) shows a clear centered distribution with very low values in large parts of the outer city. When considering the distribution of the merged overall bikeability index (bottom right), the highest values are seen in the eastern part of the city center as well as in certain other districts. The trend towards higher values in inner city districts is evident but there are dispersed characteristics as well.
Figure 4. Spatial distributions of indicators.

4.3. Influence on Mode Choice

The estimation results of the applied mode choice model presented in Section 3.3 are shown in Table 5, including the value of the estimator, the standard deviation, their respective t-values and $p$-values. The model was estimated using Biogeme [108]. The overall model performance is good with a rho square of 0.403, while all results are significant at a 5% level. Further, the parameter signs are as expected, i.e., parameter values for travel cost and travel times for all modes are negative. The age parameters indicate that the older a person is, the more likely they are to choose car (0.0254) or bike (0.0251) or even walk (0.0151), rather than use public transport. The availability of a seasonal public transport pass on the other hand reduces the likelihood of using a car ($-0.245$) compared to public transport, which is in line with expectations. Those respondents having a seasonal public transport pass have also a higher chance to bike (0.332) or walk (2.31). Taking a closer look to the travel time parameters, it can be stated that the time riding a bike is perceived much more negative ($-0.299$) than time spend on any other mode. The disutility of travel
time by public transport comes next with a parameter estimate of $-0.0324$, while the time spent walking ($-0.00271$) or driving ($-0.01$) is perceived less negative. The alternative specific constants (ASC), capturing the utility not included by other variables in the model, show a higher utility for bike and train compared to the car and lower utility for walking.

Table 5. Model results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimator</th>
<th>Standard Deviation</th>
<th>t-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC Bike</td>
<td>1.17</td>
<td>(0.0684)</td>
<td>17.1</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>ASC Car</td>
<td>0</td>
<td>(fixed)</td>
<td>fixed</td>
<td>-</td>
</tr>
<tr>
<td>ASC Train</td>
<td>1.87</td>
<td>(0.058)</td>
<td>32.4</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>ASC Walk</td>
<td>-2.03</td>
<td>(0.0537)</td>
<td>-37.8</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>Age Bike</td>
<td>0.0251</td>
<td>(0.00093)</td>
<td>27.1</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>Age Car</td>
<td>0.0254</td>
<td>(0.00091)</td>
<td>27.8</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>Age Train</td>
<td>0</td>
<td>(fixed)</td>
<td>fixed</td>
<td>-</td>
</tr>
<tr>
<td>Travel Cost</td>
<td>-0.839</td>
<td>(0.0291)</td>
<td>-28.8</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>PT Ticket Bike</td>
<td>0.332</td>
<td>(0.0391)</td>
<td>8.49</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>PT Ticket Car</td>
<td>-0.245</td>
<td>(0.0352)</td>
<td>-6.96</td>
<td>$&lt;3.38 \times 10^{-12}$</td>
</tr>
<tr>
<td>PT Ticket Train</td>
<td>0</td>
<td>(fixed)</td>
<td>fixed</td>
<td>-</td>
</tr>
<tr>
<td>PT Ticket Walk</td>
<td>2.31</td>
<td>(0.0349)</td>
<td>66.1</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>Travel Time Bike</td>
<td>-0.299</td>
<td>(0.00508)</td>
<td>-58.9</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>Travel Time Car</td>
<td>-0.01</td>
<td>(0.0033)</td>
<td>-3.03</td>
<td>0.00243</td>
</tr>
<tr>
<td>Travel Time Train</td>
<td>-0.0324</td>
<td>(0.00109)</td>
<td>-29.8</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
<tr>
<td>Travel Time Walk</td>
<td>-0.00271</td>
<td>(0.00051)</td>
<td>-5.32</td>
<td>$1.01 \times 10^{-7}$</td>
</tr>
<tr>
<td>Bikeability</td>
<td>0.518</td>
<td>(0.0623)</td>
<td>8.49</td>
<td>$&lt;2.2 \times 10^{-16}$</td>
</tr>
</tbody>
</table>

Regarding the bikeability parameter estimated by the model, a strong positive effect (0.518) of a high bikeability index along the route on choosing the bike as preferred mode can be observed. This confirms that the bikeability along a route has a significant impact on the choice of mode when making a trip. Furthermore, it shows that the bikeability index derived in this paper performs well in collecting and representing the information, which is crucial to describe bike friendliness along spatial areas.

5. Discussion

In this study, an approach to assess the bikeability of urban infrastructures using open data is demonstrated. At the same time, the relevance of individual methodological components is evaluated. These components are built on each other and form the consistent overall framework of this approach. Thereby literature review, expert survey and operationalization of parameters are crucial to establish the index while the model evaluates its applicability by quantifying significant positive influences of bikeability on mode choice.

The results of the conducted expert survey confirm the relevance of parameters determined from the literature. Key insights from the survey are the weighting itself and the consistency of the determined weighting across different subgroups of participants. This stable weighting enables an approximation of bikeability based on the joint assessment of many experts. In contrast to analyzing individual parameters [13–16], the joint integration of different parameters improves the approximation of bikeability and provides a more realistic overall picture—one main benefit of the present research. If the framework obtained in the expert survey is compared to recent integrative approaches, it is seen that the three most important categories (biking facilities along main streets, street connectivity and prevalence of neighborhood streets) are similar [21] even with some differences in
ranking and operationalizing values. Refining previous approaches [24] and basing the calculation on broad expert assessment ensures an appropriate appraisal of bikeability.

Results of geodata analysis in Berlin show that OSM open data cover each part of the city sufficiently and are accurate enough as a basis for the analysis. The method of operationalizing categories to parameters also enables an approximation that characterizes the city well with regard to bikeability and its components. In combination with high accuracy and topicality of the data [101–103], it can be assumed that this approach allows transferability and consistency between municipalities and across national borders as desired by previous studies [24]. At present, limitations on data availability still exist, particularly in Asian countries. Using Berlin as an example, it can be seen that indices for individual parameters vary significantly within the city. Each index shows a specific statistical distribution and a characteristic spatial distribution. Remaining area-based, the method is appropriate for abstracting characteristics of points (rental and repair facilities, intersections), lines (road types and coverage of biking facilities) and polygons (green areas) and for merging different parameters at any district level in urban areas.

Overall, the mode choice model confirms the influence of the bikeability index on the choice of using the bicycle as a primary mode of transport for the example of Berlin. Thereby, the model performs quite good as indicated by a rho square value of 0.403. This is especially valid against the background that choosing the bike as a mode of transport generally has several impacting factors. Especially personal attitudes and preferences are difficult to model. Thereby, 56% of the population, the “interested but concerned” are seen to be the key target market to increase cycling and, therefore, accordingly sensitive for the impacts of a bike-friendly urban environment [7]. This conversely means almost half of the population is not sensitive to differences in bikeability. Hence, modeling the usage of bike as based on spatial data is unlikely to show high coefficients of determination. Correspondingly, Winters at al. (2016) found coefficients of determination of about 0.35 [24], Dill and Carr (2003) a range from 0.18 to 0.3 [14]. Relating the results to the “interested but concerned” [7] also emphasizes the importance of the individual when measuring the impact of bikeability on bicycle usage. Prior research found a strong interrelation of objective and subjective risk when cycling even though several specific road attributes are associated with an over- or underestimation of the actual crash risk [96]. However, the areas with higher bikeability are not generally safer [99]. While these considerations are at much higher level of detail than the present index allows to render, linking bikeability and (perceived) safety is a future research topic of interest.

With a closer look, all parameters in the model perform as expected in regard to their sign and value. This is true for both cost and travel time parameters as well as additional parameters such as the availability of a public transport season ticket. The negative parameter values for cost and travel time and the more negative perception of travel time when riding a bike or using public transport compared to driving a car, which can be explained by the effort needed for long bike rides and the discomfort of travelling with other people in crowded vehicles in public transport, which are in line with other literature results [109,110].

According to the field of operation, there are various options to adopt the index: at present, different designs of cycle facilities are analyzed in an aggregated way. Here, differently recorded infrastructures (i.e., bike lane vs. bike path) could be evaluated in detail. The parameter used for road categories indicates the importance of each road for motorized transport. In the present study, this parameter approximates the level of disturbance cyclists experience due to passing vehicles. Results show that the road network configuration differs between statistic units but extreme differences in the results are rare. Future research could take account of road construction standards, such as the number of lanes or speed limits, as shown in earlier studies described in the literature review section, and could evaluate whether significant interrelations can be observed. It could also include parameters approximating bicycle parking and surface quality since these categories were mentioned in the expert survey. This information is not included in OSM
in sufficient quantity at present. Due to exponential growth in OSM, the information may be available in the future. Regarding the parameters used, this index focuses on urban infrastructure that is adaptable to a certain extent. It is, therefore, meant to understand bikeability that is in the responsibility of local or regional authorities. When aiming to draw a more comprehensive picture of bikeability than the scope of the present study, future research can easily integrate a control variable such as topography, which may also be extracted from OpenStreetMap.

Compared to earlier research, the present approach addresses shortcomings of prior studies, namely, a common definition of bikeability [9,35], a systematic valuation of relevant factors [34] and the use of open geodata [24]. From a methodological perspective, combining these issues is the main difference to valuable prior approaches, such as [17,21,24]. In addition, the present approach is fully transferable and enables the assessment of bikeability on any spatial level in large parts of the urbanized world. Hence, the present research lays the foundations for diverse large-scale analyses and may upgrade earlier results by refining the approximation of bikeability [13,21,24] and integrating comparable bikeability measures in cross-nation comparisons [12]. Regarding the results, the present study is in line with earlier studies identifying the great importance of dedicated cycle facilities [13,21,24,71] and intersection density [60,66–68].

The limitations of the approach are discussed in the following. First, there is always a trade-off between level of detail on one side and transferability and applicability on the other side. As described, in this contribution we focus on the latter. We, therefore, cannot integrate all potentially influencing factors due to limitations in data availability, differences in operationalization and different impacts for specific user groups. Second, experts may have different impression on the factors they rated. Since we aim to define a global definition, we did not predefine the factors or describe the operationalization. It is possible that some experts thought of different pictures than others. Given the large number of experts, their different background and no significant differences across groups, we believe we gathered a good mean weighting of the categories. Third, the geodata analyses limits the ways of operationalization. This is discussed in the methods section for instance regarding the parameter green pathways. Given the data and the objective of ensuring transferability, we believe that, despite uncertainties caused by operationalization issues, the developed index is a proxy for the bike-friendliness of urban infrastructures. Fourth, regarding the model, we do not know the exact routes chosen by the participants of the survey. As an approximation we use a shortest path routing from the start to the destination area. Since we assume that at least some participants accept detours to have a more pleasant routes, the bikeability of some bike trips might be higher than supposed in the model. Therefore, the influence of the bikeability index might be slightly underestimated in the model.

6. Conclusions

In the present contribution, a conceptual approach of deriving bikeability of urban infrastructure is proposed. The multifactor index is based on a literature review and an expert survey followed by geodata analyses. By developing a common weighting assessed by the group of experts, using a systematic framework including various influencing factors and analyses based on open geodata, the approach addresses shortcomings identified in prior research.

Thereby, the results of the expert survey provide a stable weighting of the influencing factors of bikeability, which are, in order of importance, biking facilities along main streets, street connectivity, the prevalence of neighborhood streets, green pathways and other cycle facilities, such as rental and repair facilities. The resulting bikeability in return shows a strong positive effect on choosing the bike as the preferred mode in the mode choice model.

Using open data only guarantees nearly unlimited transferability. Therefore, the method lays the foundation for large-scale analyses to evaluate the impact of bikeability on cycling mode shares and other dependent variables regardless of administrative borders and according to limitations in data availability. Being able to assess bikeability with high

spatial resolution in an automated process will make it possible to carry out comprehensive large-scale analyses on interrelations between bikeability and, for example, public health or collision rates. The example of Berlin shows that the index adds a significant parameter with high impact when aiming to explain choosing the bike as a mode of transport.

**Author Contributions:** Conceptualization, M.H.; methodology, M.H., M.L., S.N. and J.W.; software, S.N. and M.L.; validation, S.N. and M.L.; formal analysis, S.N., M.H., M.L. and J.W.; investigation, M.H.; data curation, M.L. and S.N.; writing—original draft preparation, M.H., S.N. and J.W.; writing—review and editing, M.H.; visualization, S.N.; supervision, M.H.; project administration, M.H.; funding acquisition, M.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** The project is funded by the German Federal Ministry of Transport and Digital Infrastructure using resources from the National Cycling Plan 2020 (NRVP).

**Data Availability Statement:** The data that support the findings of this study are available upon reasonable request from the authors.

**Acknowledgments:** The authors acknowledge all experts who participated in the survey.

**Conflicts of Interest:** The authors declare no conflict of interests.

**References**


45. Wang, Y.; Chau, C.K.; Ng, W.Y.; Leung, T.M. A review on the effects of physical built environment attributes on enhancing walking and cycling activity levels within residential neighborhoods. *Cities* **2016**, *50*, 1–15. [CrossRef]


64. Badland, H.M.; Schofield, G.M.; Garrett, N. Travel behavior and objectively measured urban design variables: Associations for adults traveling to work. Health Place 2008, 14, 85–95. [CrossRef] [PubMed]
65. Fan, J.X.; Wen, M.; Kowaleski-Jones, L. An ecological analysis of environmental correlates of active commuting in urban U.S. Health Place 2014, 30, 242–250. [CrossRef]
70. Bopp, M.; Gayah, V.; Campbell, M. Examining the Link Between Public Transit Use and Active Commuting. Int. J. Environ. Public Health 2015, 12, 4256. [CrossRef]
75. Parkin, J.; Wardman, M.; Page, M. Estimation of the determinants of bicycle mode share for the journey to work using census data. Transportation 2008, 35, 93–109. [CrossRef]


