Accident Analysis and Prevention 121 (2018) 109–117

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Risk perception during urban cycling: An assessment of crowdsourced and authoritative data

Rul von Stülpnagel[a,*], Jakub Krukar[b]

[a] Center for Cognitive Science, Freiburg University, Hebelstr. 10, 79104, Freiburg, Germany
[b] Institute for Geoinformatics, Münster University, Heisenbergstraße 2, 48149. Münster, Germany

ARTICLE INFO

Keywords:
Urban cycling
Risk perception
Volunteered geographic information
Space syntax
Accident statistics

ABSTRACT

Subjective risk perception during urban cycling has been mostly investigated through questionnaire studies. However, newly available data sources promise extended possibilities for the investigation and understanding of the underlying factors. We validate the rationale for using both opportunistically available crowd-sourced data (i.e., volunteered geographic information or VGI) as well as more established but rarely investigated authoritative data as predictors of subjective cycling risk. We achieve this by correlating indicators of cycling risk extracted from both VGI and authoritative data for two different German cities with participants' risk estimates assessed in laboratory-based virtual reality experiments. In Case 1, 15 participants (mostly undergraduate students with a mean age of 22 years old; nine of them females) were tested as a sample representing frequent and experienced cyclists, but unfamiliar with the 19 tested locations and less likely to be affected by the virtual reality setup. In Case 2, 24 new participants (mostly undergraduate students; mean age 24 years; 13 of them females) were experienced cyclists and mostly familiar with the 40 test locations located in their city of residence. For both cases, our findings provide evidence that parameters extracted from VGI (e.g., the semantic severity of the contribution and the reception by other citizens) as well as from authoritative data sources (e.g., accident statistics or Space Syntax measures) represent valid indicators for the subjectively perceived risk of cycling at a specific location. On the basis of this validation, future research can use these data sources to investigate the sources of risk perception during urban cycling in greater detail.

1. Introduction

Cycling is a healthy and environmentally friendly alternative to driving and to the use of public transport for urban commuting. Consequently, several governmental and municipal programs aimed at increasing the popularity of cycling (e.g., "Encouraging Cycling in Central London | Space Syntax," 2000). Simultaneously, a number of studies examined factors preventing people from using their bike (see Fernández-Heredia et al., 2014, for an overview). As one major deterrent to cycling, Wardman et al. (1997) identified the associated risk and unpleasantness of typical urban traffic conditions to be more discouraging than the effort of cycling itself. However, several researchers also pointed out that subjectively perceived hazards remain an understudied subject with a considerable impact on cycling volume (Chaurand and Delhomme, 2013; Rietveld and Daniel, 2004; Schepers et al., 2014; Sørensen and Mosslemi, 2009).

The majority of studies on this matter investigated subjective risk perception during cycling via questionnaires. Parkin et al. (2007) assessed the perceived cycling risk for complete routes. Chaurand and Delhomme (2013) instructed participants to imagine themselves in several proth2otypical traffic situations and to rate their probability of being involved in an accident within the next three years. Chataway et al. (2014) assessed subjective fear of traffic with items referring to, for example, cars passing in close proximity to the cyclist. Additionally, they presented abstract illustrations of traffic situations that the participants rated from "unsafe" to "safe". Another study collected a large body of questionnaire data in order to investigate the relation between self-reported accidents and perceived risks of cycling (Washington et al., 2012). It can be argued that investigations of subjective risk during cycling are limited by the amount and representativeness of data that can be collected via questionnaires. In this research, we thus aim to assess the potential of two alternative sources on subjective cycling risk, namely crowd-sourced data and authoritative information sources.

* Corresponding author.
E-mail addresses: rul.von.stuelpnagel@cognition.uni-freiburg.de (R. von Stülpnagel), krukar@uni-muenster.de (J. Krukar).

https://doi.org/10.1016/j.aap.2018.09.009
Received 7 May 2018; Received in revised form 11 September 2018; Accepted 12 September 2018

1.1. Assessing subjective cycling risks from crowdsourced data

The development and widespread availability of internet-based crowd-sourcing tools have opened up new possibilities for the collection of large datasets from the general population. Platforms sponsored by non-governmental organizations (NGOs) such as the UK-based Collideoscope invite citizens to contribute information about obstacles for urban cycling on their websites. Academia-based approaches such as the BikeMaps project allow contributors to enter information on cycling collisions, near misses, and various hazards in a semi-structured format (Nelson et al., 2015). Tapping into citizens' reports promises to be an unprecedented and valuable information source. However, this promise comes with a number of challenges attached: the majority of these projects allow contributions in an open text format, and citizens' reports may thus consist of idiosyncratic and fuzzy descriptions of anticipated (rather than actually experienced) risks when cycling at a given location. It is currently unknown whether the provided information is of the necessary quality and representativeness to be used in investigations of subjective risk perception during urban cycling.

Crowdsourced datasets consisting of citizens' contributions at explicit geographic locations have seen increased attention under the label of *volunteered geographic information* (VGI; Goodchild, 2007). Consistent with the idea of 'collective intelligence' (Spielman, 2014), researchers in this field argue that VGI carries the potential for providing information superior to any single existing source of the same information. The rapidly growing popularity of VGI has also generated a discussion on its usefulness and quality (e.g., Elwood et al., 2012). Goodchild and Li (2012) review three ways to assure the quality of VGI: crowdsourcing (correction of errors and mistakes by the crowd itself), social (through a hierarchy of administrators and gatekeepers), and geographic (by possibly automated matching with 'true' properties of space, such as its geometry). However, this classification is mostly limited to assuring the *accuracy* of VGI, that is, to cases where it is possible to converge on an objective 'truth' of the topic in question (e.g., a building's exact geographic location; Haklay, 2010). Another defined the "seven pillars" of VGI quality (e.g., LBS-Positioning, Authoritative Data Comparison, and Semantic Harmonisation; Leibovici et al., 2017; Meek et al., 2014).

In the particular case of risk perception during urban cycling, however, the provided information is not an objectively measurable physical property of the world, but the opinion of the individuals inhabiting it: The subjective impression of risk expressed in such a contribution remains true even if it is not matched by a detectable accident risk. Thus, approaches suited to estimate the accuracy of VGI cannot be easily extended to determine the validity of subjective risk described in a VGI contribution. In this research, we take a closer look at two parameters potentially providing information in this regard that can be extracted from a VGI dataset alone. First, the content of a VGI contribution itself, and second, the reception of a VGI contribution by others.

1.1.1. Content analysis

The setup of many public VGI projects concerning cycling risks is not guided by scientific considerations, and the data acquisition for scientific purposes is often opportunistic (but see the BikeMaps project by Nelson et al. (2015), for a more structured approach). Consequently, contributions are rarely standardized and frequently consist of an open text format encompassing an underspecified topic. The challenge is to standardize and compare such input. Within the framework of the seven pillars of VGI quality, this challenge falls into the category of "semantic harmonization" (Leibovici et al., 2017; Meek et al., 2014). Several approaches addressing this matter used term frequencies or semantic similarities (for a review, see Schwering, 2008). These approaches, however, are more geared towards enhancing the compatibility of contributions with different terminologies and from different sources. They are less suitable to account for VGI reflecting subjective impressions containing emotional connotations with large inter-individual differences. Finally, some studies aimed to integrate this emotional level into the analysis of GIS data (e.g., Abdalla and Weiser, 2011; Hauthal and Burghardt, 2013). However, these approaches were more concerned with the spatial distribution of emotional connotations, and less with the emotional appraisal in the description of a VGI contribution.

1.1.2. Public reception

Another approach for determining the level of subjective risk during urban cycling mentioned in a VGI contribution consists of examining the endorsement by others. The underlying assumption is that the more people experience subjective risk at a given location, the more likely they are to indicate agreement with an existing contribution. Functions such as Facebook's 'Like' button may represent a way to tap into the relevance and representativeness of the original contribution (Jessen and Jørgensen, 2011). Bishr and Kuhn (2007) proposed to utilize user trust as a proxy for VGI quality (see also Leibovici et al., 2017; Meek et al., 2014). This approach cannot easily be applied directly to the raw data as investigated in this research, as such contributions may be distorted by the contributors' uneven activity on the platform (Dubrovsky et al., 1991).

1.2. Links between subjective risk perception and authoritative data

Authoritative data have been previously used to assess cycling risks. In this research, we consider two authoritative data sources; namely accident statistics and traffic infrastructure, the latter operationalized through Space Syntax measures.

1.2.1. Accident statistics

One authoritative data sources that can be specifically linked to risk perception during urban cycling is the (frequently believed to be objective) accident statistics (Gregoriades and Chrystodoulides, 2018; Molino et al., 2009). However, accident statistics must be assumed a problematic and potentially confounded indicator: Perceived risk during cycling does not necessarily influence injury rates, nor do injury rates influence perceived risks of cycling (Washington et al., 2012). This may partly stem from the fact that cyclists avoid or take extra care in areas they believe to be problematic (Parkin et al., 2007). Furthermore, accident statistics do not account for the comparatively much larger number of "near misses" (i.e., events where a car passes dangerously close to the cyclist, Sanders, 2015), although recent findings show a similar distribution of near misses and real accidents, suggesting that they can be used to complement gaps in authoritative datasets (Branion-Calles et al., 2017). Taken together, the rather heterogeneous findings highlight the need for further investigations of the relation between perceived cycling risks and accident statistics.

1.2.2. Traffic infrastructure

Another source of authoritative data potentially relevant for subjective cycling risk is the traffic infrastructure at a VGI contribution's geolocation. A larger street size (and thus a higher traffic volume), as well as more complex traffic situations, have been demonstrated to increase mental workload (Jahn et al., 2005), which in turn increases the risk of accidents (Brookhuis and de Waard, 2001). It can be assumed that such locations are perceived as consistently more dangerous.

The complexity of a location's traffic infrastructure can be approximated through *Space Syntax* (Hillier, 1996; Hillier and Hanson, 1984), which utilizes graph theory to study the configurational relations between the interconnected elements of a street network in a so-called axial map analysis. In an axial map, the street network is represented as a set of straight lines, each corresponding to the longest available line of sight (i.e., a single curvy street might be thus represented by more than one axial line). Each axial line can be assigned various properties depending on its position in the overall graph, for example, the number of connections with other lines (for an introduction, see Bafna, 2003). *Integration* is a standard measure in axial maps reflecting how easy it is to access all other streets in the city from the street at hand. The integration value of a given axial line has been found tightly linked to the size and traffic volume of the street it represents (Hillier et al., 1993). We thus hypothesize that a street should be perceived as more dangerous, the higher the respective integration value.

1.3. Aim of the present research

Our aim is to validate the rationale for using both crowd-sourced and authoritative data as predictors of subjective cycling risk. The availability of crowdsourced data for scientific purposes in particular is frequently opportunistic. We thus tackle our aim by applying the same experimental design to two such datasets from two different cities: Participants studied selected sets of test locations in a controlled lab-based setup. In order to keep the visual impression as constant as possible while providing a naturalistic and immersive impression at the same time, participants wore a head-mounted display (i.e., an Oculus Rift) to study the scenes. We correlate the participants' risk estimates of cycling at the presented locations from the lab with parameters extracted from both crowd-sourced and authoritative data of the datasets. If we find significant, positive relations between the lab-based estimates and the properties extracted from VGI and authoritative data, this would support the conclusion that these data are indeed valuable sources for further investigations of subjective cycling risks.

We emphasize that it is not our aim to investigate which factors initially cause risk perception during urban cycling, nor to identify the best among several indicators, but rather to establish the foundations potentially allowing for such an investigation.

2. Method

2.1. Description of the investigated datasets

Data from two unrelated crowdsourcing projects addressing urban cycling risks were available and investigated in the research at hand.

2.1.1. Case 1: Munich - the "Munich Hazard Atlas" by Sueddeutsche Zeitung

Case 1 was based on data collected in a web-based inquiry of the major, Munich-based German newspaper "Sueddeutsche Zeitung". The newspaper launched an interactive map (the so-called "Gefahrenatlas" or "Hazard Atlas") in July 2014 and invited their readers to mark locations in the greater area of Munich they considered (based on their personal experiences and observations) to be dangerous, confusing, or a nuisance in any other way for cyclists[1]. An open description of the reported hazard could be added to the marking. Existing reports of other readers could be supported via a "Like" function. The inquiry generated 5000+ unique contributions featuring open-text description, their exact geolocations, and 17,000+ Likes of contributions by other readers (see top panel of Fig. 1 for an excerpt).

2.1.2. Case 2: Freiburg - the initiative "Better on The Way in Freiburg"

Case 2 was based on a dataset collected by the NGO "*Besser unterwegs in Freiburg*" ("Better on the way in Freiburg") with about 1000 individual contributions as of summer 2015 (see Fig. 2)[2]. Similarly to the dataset investigated in Case 1, citizens were invited to contribute open descriptions about hazards for cyclists in Freiburg city[3]. The platform "*Besser unterwegs in Freiburg*" does not feature a "Like" function, but allows open format responses to the original contribution. With about 680 responses in the complete dataset, this option was much less used than the "Like" function available in the Munich project.

In addition to the VGI dataset, information about accidents involving cyclists in Freiburg was provided by the Freiburg police department. This dataset included about 1700 incidents that took place between 2012 and 2014. It included their geolocation, timestamp, the number of injured persons, and accident type (according to the German police accident classification scheme).

2.1.3. Comparison of the two investigated cities

Whereas the structure of the two projects collecting citizens' reports about cycling risks was fairly similar, the cities targeted by these projects were not: Munich is the third-largest German city with about 1.5 million citizens, covering an area of about 300 km^2. As of 2011, cycling accounts for about 17% of all traffic in Munich. Freiburg has a population of about 225,000 people and covers an area of about 153km^2. Amounting to 34% of all traffic in Freiburg in 2016, cycling accounted for one of the highest proportions in German cities. It could be argued that cycling risks are perceived and reported differently due to the cities' different sizes and traffic infrastructure. However, it is not the aim of this research to provide a detailed explanation of the sources of subjective cycling risks, but rather to provide the means to investigate them. If we find that indicators of subjective cycling risk are representative to a larger population tested in the lab despite the mentioned differences, this would highlight the relevance of the investigated data sources.

2.2. Analysis of the datasets

2.2.1. Case 1: Munich

We selected 19 contributions from the Munich dataset based on the following requirements: (1) *General area and topic*: contributions had to refer to cycling-related hazards in central Munich. (2) *Completeness of the Street View image*: Participants studied Google Street View images (see 2.4.1). Google Street View blurs faces, license plates, and some buildings to accommodate privacy concerns. In order to ensure a sufficient level of immersion, contributions were only included if their corresponding Street View image did not consist of large blurred Sections (3) *Proximity of the VGI contribution and the corresponding Google Street View image*: We selected VGI contributions with the smallest possible distance to the next available Google Street View image ($M = 7.84$ m, $SD = 6.23$).

We extracted three potential indicators for subjective risk perception:

(1) *Content Analysis*. Based solely on the content of a given VGI contribution, its *semantic severity* was evaluated by four independent raters, who were not familiar with the selected locations or Munich city in general. They read the open format comments and rated their severity of the described hazard on 7-point scales (with higher values representing a higher severity), without studying images of the actual location or knowing the number of votes. For each VGI contribution, we computed the average of the four ratings (range: 1.83-6.00, $M = 3.99$, SD=.99).
(2) *Public Reception*. We extracted the number of votes that each contribution received from other readers (range: 2-90, $M = 38.26$, $SD = 24.13$).
(3) *Authoritative Data–Traffic Infrastructure*. We used an axial line map of Munich provided by Space Syntax Ltd. and the Space Syntax Toolkit plugin for QGIS to match each VGI contribution to the integration value of its nearest street (range: .94-1.56, $M = 1.33$, SD=.18). For an illustration and further information, see Fig. 1 and Section 1.2.2, respectively. For technical details of this procedure, see the Space Syntax methodology handbook (Al-Sayed et al., 2014).

2.2.2. Case 2: Freiburg

We selected 40 VGI contributions at spatially distant locations with

[1] http://problemstrassen.sueddeutsche.de. The initiative accounted for different travel modes (i.e., pedestrians, cyclists, or skaters). Only reports concerning cycling hazards were included in the present research.

[2] http://www.besser-unterwegs-in-freiburg.de.

[3] *Besser unterwegs in Freiburg* also allows for contributions concerning "comfort" issues, as well as hazards reported by pedestrians. Contributions of these categories were not considered in the present research.

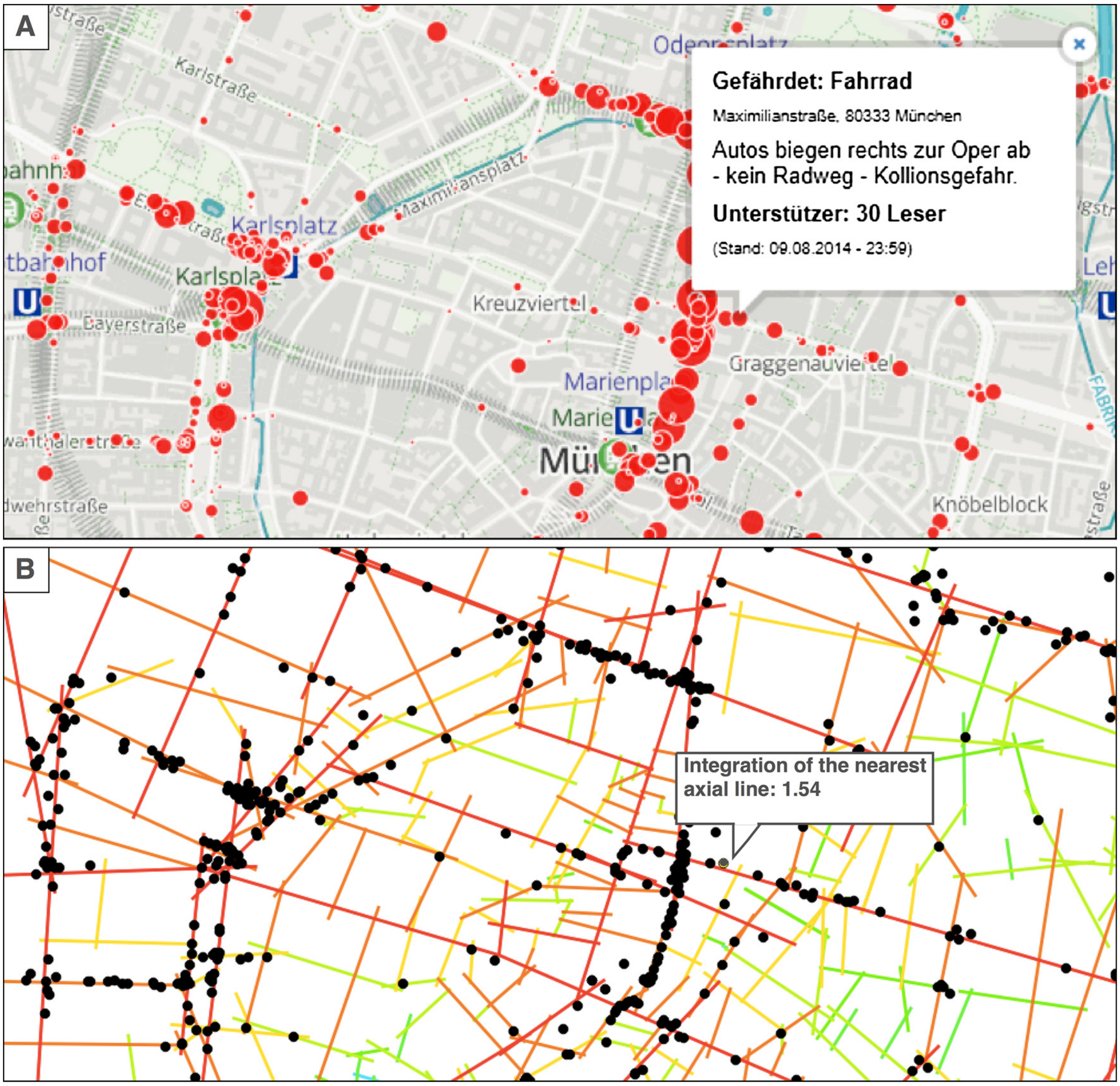

Fig. 1. *A (top panel)*: Excerpt from the VGI dataset collected by Sueddeutsche Zeitung (Case 1: Munich). Round dots (red in the web version of this article) mark individual contributions, with their size indicating the number of votes by other readers. The white box provides information about an exemplary contribution. The provided hazard description reads *"Cars turn right towards the opera house – no cycling track – risk of collision"*, with 30 votes ("Unterstützer") supporting this statement. Map excerpt: © Sueddeutsche Zeitung. *B (bottom panel)*: A corresponding fragment of the axial line map using the natural breaks coloring scheme. Darker lines (red in the web version of this article) indicate higher integration; lighter lines (green in the web version of this article) indicate lower integration. Contributions were linked to the nearest axial line, resulting in an integration value associated with each contribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the following considerations: (1) They were described from the cyclist's perspective. (2) They referred to hazards due to cars taking a turn and potentially overlooking cyclists. (3) They were located either within 10 m to a known accident location (as indicated by the accident statistics provided by the Freiburg police department, or further than 30 m away from the nearest known accident location (see (3) *Authoritative Data - Accident Statistics* below).

We extracted the three following indicators:

(1) *Content Analysis*. Corresponding to Case 1, sixteen independent raters estimated the level of *semantic severity* of each contribution on a 7-point scale (*"Imagine yourself in a cyclist's position. Please estimate the severity of the described situation?"*; 1 = not dangerous at all; 7 = severely dangerous), based solely on wording and connotation of each VGI reports (i.e., without visual information about the respective location). In order to avoid that the raters' estimates of the semantic severity are confounded by their individual memories and experiences at these locations, we substituted all proper location names with the words "sun", "moon", and "star" (e.g., the statement "Cyclists travelling on Main street are frequently overlooked by drivers turning into Tower street" would be changed to "Cyclists travelling on Sun street are frequently overlooked by drivers turning into Moon street"). We computed the average of these ratings per location (overall $M = 4.81$, $SD = .51$).

(2) *Public Reception*. We extracted the *number of responses* for the 40 selected contributions (range: 0-8, $M = .90$, $SD = 1.46$).

(3) *Authoritative Data - Accident Statistics*. The 40 selected VGI contributions fell into two categories. For twenty VGI contributions, the accident statistics indicated no accident within a 30 m radius (*Accident category: VGI only*). For the other 20 VGI contributions (*Accident category: VGI + accident*), at least one accident of the accident categories *turning accidents* and *crossing accidents* (*range*: 1–6, $M = 2.30$, $SD = 1.75$) had occurred within a 10 m radius. Both categories refer to situations of drivers overlooking cyclists when taking a turn. For an illustration, see Fig. 2.

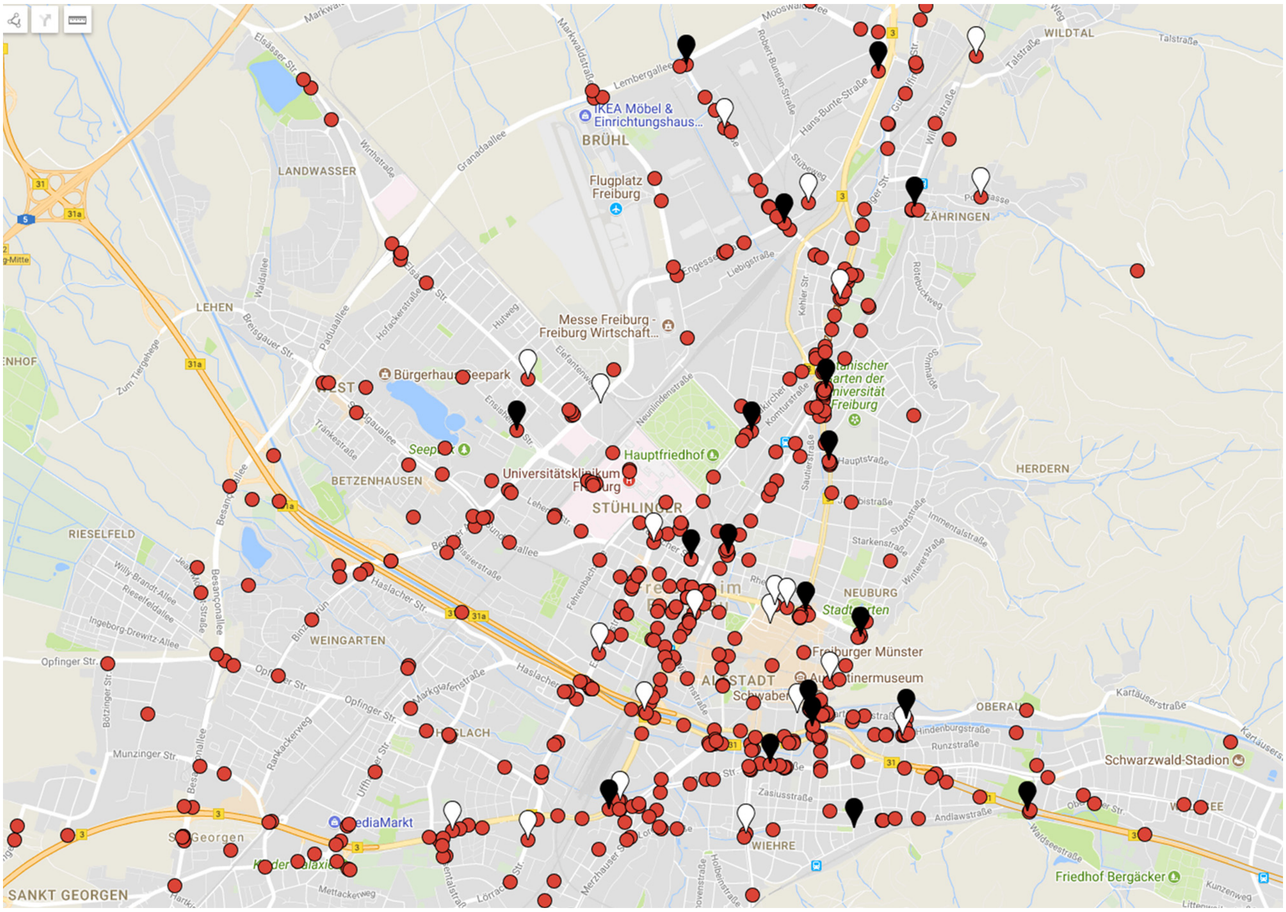

Fig. 2. Illustration of the VGI project "*Besser unterwegs in Freiburg*" (Case 2: Freiburg). Round dots (red in the web version of this article)indicate individual contributions (with multiple contributions overlapping in the city center). White pins indicate the twenty test locations of the accident category "VGI only". Black pins indicate the twenty test locations of the accident category "VGI + accident". Map excerpt: © Google Maps. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.3. Participants

2.3.1. Ethics statement

All participants signed informed consent at the beginning of the experiment. There was no foreseeable harm from participation in the study; participants were not deceived; all collected data was strictly anonymized. Participants could abandon the study at any stage without providing a reason.

2.3.2. Case 1: Munich

We tested 15 students of Freiburg University (mostly undergraduates from various course programs; nine females; age: 19–26 years, $M = 22$, $SD = 2$), who participated in exchange for course credit or €8. Testing undergraduate students living in Freiburg ensured a participant sample with sufficient and recent cycling experiences. Such a sample is likely to provide valid estimates of the presented scenes without experiencing irritations from the technical apparatus.

As expected, their self-reports indicated that the presentation of the test locations via the Oculus Rift was perceived as quite natural (*"How much did your experiences in the virtual environment seem consistent with your real-world experiences?"*; $M = 5.00$, $SD = 1.36$ on an 8-point scale) and not affected by cybersickness (Experiences of general illness: $M = 1.79$, $SD = 1.31$ on a 6-point scale). Participants estimated their cycling experience to be high (*"How do you estimate your cycling ability?"*; $M = 6.07$, $SD = 1.38$ on a 7-point scale as all following items) and frequent cycling activities (*"How often do you cycle in the city center?"*; $M = 5.43$, $SD = 2.47$). Participants perceived urban cycling as a rather dangerous activity (*"How dangerous do you estimate cycling in urban traffic?"*; $M = 4.86$, $SD = .95$) and were unfamiliar with Munich (*"How well do you know the city center of Munich?"*; $M = 1.86$, $SD = 1.03$).

2.3.3. Case 2: Freiburg

Participants were selected with a reasoning corresponding to Study 1. Twenty-eight students and inhabitants of Freiburg (13 females; age: 18–33 years, $M = 24$, $SD = 4$) participated in exchange for course credit or €8. The enhanced presentation format (see 2.4.2) was perceived as quite natural ($M = 6.36$, $SD = 1.13$) without inducing cybersickness ($M = 2.46$, $SD = 1.79$, both items assessed on 8-point scales). Again, self-reports (all assessed with 7-point scales corresponding to those described in 2.3.2) indicated high cycling experience ($M = 5.89$, $SD = 1.10$) and frequent cycling activities ($M = 5.14$, $SD = 1.86$). Urban cycling was perceived as a rather dangerous activity ($M = 5.14$, $SD = 1.51$). Familiarity was assessed separately for each test location ($M = 3.97$, $SD = 2.24$ across all participants and test locations).

2.3.4. Participants' familiarity with the test locations

The familiarity of the participants tested in the lab with the presented locations differed for the two investigated cases. In Case 1, participants recruited at Freiburg University were unfamiliar with the actual locations located in Munich. Thus, their lab-based hazard perception should be only informed by the ad-hoc visual evaluation of the immediate surroundings and the general experience with urban cycling, but not by previous exposure or knowledge about local specificities. It can be questioned to what extent the participants' perception of an unknown traffic situation derived from a single, static image is valid, or whether they are able to estimate cycling risks at all. The investigation of Case 2 can shed light on this matter, as participants were mostly familiar with the actual locations.

2.4. Materials & procedure

2.4.1. Case 1: Munich

All participants were tested individually on a swivel chair. They were told that they were to study and rate several street scenes from a cyclist's perspective. Participants were exposed to the selected Google Street View locations by wearing an Oculus Rift DK2 head-mounted display with a resolution of 1920 × 1080 pixels and a horizontal field of view of about 90°, which allowed them to study a test location from one specific point of view while looking around in a natural fashion. The experimenter asked them to close their eyes until the respective image was loaded. Participants were told to imagine cycling at the presented locations and to study the scene until they had formed an impression about the risk of cycling at this location (*Lab-estimate*), which they rated on a 7-point scale (1 = not dangerous at all; 7 = severely dangerous). This procedure was repeated for all 19 locations in a randomized order. The first presented scene was used by the experimenter for adjusting the head-mounted display and familiarizing the participants with the procedure, and thus excluded from the analysis. Finally, information about demographic data, cycling experience, and familiarity with Munich central area was assessed with a questionnaire before participants were thanked and debriefed. The whole experimented lasted about one hour.

2.4.2. Case 2: Freiburg

The procedure resembled the one described above with the following exceptions. Google Street View is not available for Freiburg city. Instead, we used the photosphere mode of an Android-based smartphone to create 360° high fidelity panosphere images with a resolution of 9728 × 4864 pixels for all 40 locations as indicated by the respective geo-reference on www.besser-unterwegs-in-freiburg.de. Each participant studied a pseudo-randomized subset consisting of 20 locations, with the aim that each location was presented to about fifteen participants. Next to estimating the risk of cycling at a presented location (*Lab-estimate*), participants also rated their *familiarity* with each location on a 7-point scale (see Fig. 3).

3. Results

We aimed at evaluating each location's subjective level of danger by linking several indicators extracted from VGI and authoritative data to the baseline of participants' lab-based risk estimates. Statistically, this comparison can be presented as a linear model, where the laboratory-based estimate is the dependent variable (or the 'output'), while the available parameters extracted from VGI and authoritative data are used as independent variables (i.e., 'predictors'). We constructed two separate linear mixed-effect model to express this in a statistical analysis without aggregating the results by-participant or by-location.

All analyses were conducted with the software Statistical Package for the Social Sciences (SPSS), version 24.

3.1. Case 1: Munich

With 15 participants contributing estimates for 18 test locations each, 268 data points were available (with 2 estimates missing due to technical issues).

Including all parameters (i.e., a VGI's semantic severity, its number of votes, as well as the integration value of the closest street) as predictors of the lab-bad risk estimate resulted in the following model:

$$Lab\ estimate\ \sim\ Semantic\ severity + Number\ of\ votes + Integration + (1|participant\ .\ ID)$$

Both the semantic severity of a cycling hazard described in a VGI contribution as well as the public response to such a contribution (as operationalized by the number of votes from other citizens) showed a significant relation to the estimates of participants in the lab (see Table 1 for fixed effects and Table 2 for parameter estimates). Thus, our results demonstrate that information extracted from non-scientific, spontaneously contributed VGI is a relevant source for assessing subjective risk perception during urban cycling. Additionally, our results imply that participants were sensitive to street size and complexity (operationalized with a Space Syntax analysis), as more integrated locations were perceived as more dangerous.

3.2. Case 2: Freiburg

With 28 participants contributing estimates for a subset of 20 of the 40 test locations, 560 data points were available.

We constructed a linear mixed-effect model with *Semantic severity*, *Number of responses*, and participants' *familiarity* with the individual location as continuous predictors of the dependent variable *Lab estimate*. We included *accident category* as a categorical variable with the two levels *VGI only* and *VGI + accident*. This resulted in the following model:

$$Lab\ estimate\ \sim\ Semantic\ severity + Number\ of\ responses + VGI\ density + Familiarity\ .\ participant\ .\ ID + Accident\ category + (1|participant\ .\ ID)$$

Familiarity proved to be a highly significant predictor: the more knowledge about a location, the more dangerous it was estimated to cycle there (see Table 3 for fixed effects and Table 4 for parameter estimates). Despite the participants' greater familiarity with the tested locations, semantic severity remained a significant predictor of risk perception. The number of responses showed a nearly significant trend in the same direction (p = .06). We reason that - in comparison to the number of votes investigated in Case 1 - the open format responses of the dataset used in Case 2 was not only less frequently used, but also allowed for more diverse content (e.g., a response could deny the existence of the hazard described in the original contribution, yet it would still increase the number of responses).

Concerning our analysis of accident statistics as an instance of authoritative data, our model indicates that test locations with no previous accidents were perceived as significantly less dangerous as compared to locations with one or more previous accidents. Thus, participants in the lab (who were cycling frequently in Freiburg and were mostly familiar with the test locations) are apparently sensitive to the increased accident probability.

4. General discussion

The perception of traffic risks represents one of the major deterrents for cycling (e.g., Fernández-Heredia et al., 2014; Wardman et al., 1997). The majority of studies on this matter investigated subjective risks via questionnaires (e.g., Chataway et al., 2014; Chaurand and Delhomme, 2013; Parkin et al., 2007). In the present research, we aimed at extending these approaches by assessing crowdsourced data (i.e., VGI) as well as authoritative data as potential information sources indicating subjective risk perception during urban cycling. Participants studied images of several real-world locations and rated the estimated risk of cycling at these locations in two lab-based studies. This baseline estimate was compared with several parameters extracted from the investigated datasets. Our results support the assumption that both VGI and authoritative data provide valid information about risk perception during cycling, irrespective of differences between the two investigated cities' and the participants' differing levels of familiarity with the tested locations. Our findings are now discussed in greater detail.

Lab Estimates	
Risk Estimation	3.86
Familiarity	3.71
VGI Properties	
Semantic Severity	4.69
Number of Responses	3
Authoritative Data	
Number of Accidents	1

Fig. 3. Illustration of a panosphere image presented via the Oculus Rift in Case 2: Freiburg. The respective VGI contribution states: *"Most car drivers ignore the right of way of people coming from Mühlenweg and turning right, the more so as the view into Mühlenweg is poor. A warning sign on Karthäuserstrasse to respect the right of way would help."* The table on the right presents all the analyzed properties of this respective location.

Table 1
Tests of fixed effects for the linear mixed model in Study 1.

Source	Numerator df	Denominator df	F
Intercept	1	256.55	.09
Semantic Severity	1	248.29	12.90**
Number Of Votes	1	247.45	4.65*
Integration	1	248.60	4.53*

* $p < .05$.
** $p < .001$.

Table 2
Parameter estimates of fixed effects for the linear mixed model in Study 1.

Parameter	*B*	SE	df	t
Intercept	−.33	1.09	256.55	−.30
Semantic Severity	.41	.12	248.29	3.59**
Number Of Votes	.01	.00	247.45	2.16*
Integration	1.36	.64	248.60	2.13*

* $p < .05$.
** $p < .001$.

Table 3
Tests of fixed effects for the linear mixed model in Study 2.

Source	Numerator df	Denominator df	F
Intercept	1	556.93	12.05**
Semantic Severity	1	532.10	6.47*
Number Of Responses	1	532.02	3.62
Accident Category	1	531.99	7.92*
Familiarity.participant.ID	1	554.56	35.57**

* $p < .01$.
** $p < .001$.

Table 4
Parameter estimates of fixed effects for the linear mixed model in Study 2.

Parameter	*B*	SE	df	t
Intercept	2.13	.59	558.34	3.62**
Semantic Severity	.28	.11	532.10	2.54*
Number Of Responses	.07	.03	532.02	1.90
Accident Category (VGI only)	−.31	.11	531.99	−2.81*
Familiarity.participant.ID	.17	.03	554.58	5.96**

* $p < .01$.
** $p < .001$.

4.1. Risk perception and VGI

Our research displays one of the first attempts to assess open-format VGI as an information source about subjective risk perception. Web-based crowdsourcing platforms collecting geographic information provide new and already widely used means to collect data for a broad range of topics. However, the quality of VGI remains a disputed issue (Elwood et al., 2012). In the particular case of subjective risk perception, VGI cannot be easily matched against a "true" physical entity: The perception of risk remains valid even in the absence of a traceable hazard. Additional challenges result from the open text format of the investigated datasets. In this research, we presented two parameters derived from VGI that proved to be significant predictors of risk perception assessed in the lab, namely a contribution's semantic severity and its public reception.

A contribution's semantic severity appears to represent a highly informative measure concerning subjective risk perception. However, extracting the semantic severity from VGI requires considerable effort, which makes this measure impractical for large datasets. One possible solution are automated text mining approaches (Schwering, 2008), but the diversity and frequent vagueness of VGI contributions require further technical advances. Ideally, projects on cycling risks should implement the option of ranking the severity of a contributed problem directly into their platform (e.g., the BikeMaps project by Nelson et al., 2015) to circumvent this challenge.

Our research also demonstrates that more accessible parameters derived from VGI can be informative on the location's level of subjective cycling risk. The 'Like'-function investigated in Case 1: Munich is of particular interest. However, the distribution of such a function is likely to grow exponentially rather than linearly for both spatial (Jiang, 2015) and social (Cialdini, 2006) reasons. Additionally, the phenomenon known as 'participation inequality' not only implies that very few users may be responsible for the majority of the contributions, but also that an existing contribution can affect further contributions (Dubrovsky et al., 1991). For the datasets investigated in this research, we can only speculate to what extent a small number of influential opinion leaders may have affected both the distribution and the number of votes. However, the comparison of the high number of "Likes" in Case 1 with the comparatively low number of responses in Case 2 implies that features with a low participation threshold may help to mitigate effects of participation inequality in crowdsourcing projects, and thus generate valuable information about a contribution's representativeness (Bishr and Kuhn, 2007; Jessen and Jørgensen, 2011).

The raw number of open format responses as investigated in Case 2: Freiburg appears to be a parameter less suited for such an analysis due to the lower overall number as well as a more diverse content. An indepth analysis of the responses' content was beyond the scope and the available dataset of the research at hand. Despite these limitations, the raw number of responses was marginally significant as a predictor of the lab-based risk estimates in Case 2: Freiburg.

4.2. Risk perception and authoritative data

Authoritative data sources will remain relevant for urban planners and researchers alike, as they provide an established information with sufficiently well-understood potential and limitations. However, attempts to link accident statistics with risk perception are not converging on a clear pattern so far (see Gregoriades and Chrystodoulides, 2018; Molino et al., 2009; Washington et al., 2012). For Case 2: Freiburg, we found that locations where accidents had previously occurred were estimated to be more dangerous than accident-free locations. This finding implies that cyclists familiar with the city are mostly aware of potentially dangerous locations. It also suggests that accident statistics may provide not only information about the "objective" risk but also some insights concerning subjective risk perception. However, this relation is unlikely to be linear and the matter deserves further investigation: Cyclists who believe a certain location to be dangerous will adjust their behavior accordingly, which in turn may reduce the number of actual accidents (Parkin et al., 2007). A recent study featuring mobile eye tracking also provided preliminary evidence that cyclists gaze at areas where they felt an accident could occur rather than at areas where accidents actually had previously occurred (Schmidt and von Stülpnagel, 2018a,b).

As another instance of authoritative data, we tested the relation between risk perception and Integration, a standard measure of the Space Syntax approach in Case 1: Munich. Space Syntax has already been used to estimate cycling traffic volume (e.g., "Encouraging Cycling in Central London | Space Syntax," 2000; Raford et al., 2007). We found that the integration value of a street (i.e., the relative centrality of this street within the network of the city, and thus an indicator of street size, traffic volume, and complexity) contributed significantly to the level of the perceived risk. Again, this finding must be seen as initial evidence that Space Syntax measures can provide insights about subjective risk perception in general, which can be more closely investigated in future studies.

4.3. Limitations

Challenges tied to the different investigated parameters were discussed in the sections above. However, our approach also comes with some methodological limitations. First, we selected a small sample of test locations from two datasets. In addition to the constraints reported in Section 2.2, this selection was limited, for example, by the quality and location of suitable Google Street View images in Case 1. Thus, we cannot guarantee the representativeness of the test locations for the entire datasets although the selection within these constraints was random. Furthermore, the number of testable locations was limited by the time and resources required to test participants in the lab.

Second, we used an Oculus Rift to present the test locations in a controlled but immersive way. The participants' reports confirm our goal to provide an immersive and naturalistic impression of the tested locations. Nevertheless, the presented image was a static view providing a single perspective on the recording location. Possible approaches to enhance the presentation format are reported in the following section.

4.4. Future directions

Based on this research, it is now possible to use VGI to investigate subjective risk perception on a city-wide level. This allows for enhanced analyses of the factors causing subjective risk, and their relation with accident statistics and infrastructure. Such an investigation needs to account for the volume of cars and cyclists traveling at a location, and therefore the potential number of people experiencing, reporting or confirming cycling risks. This could be achieved through an extended Space Syntax analysis, which has been shown to predict up to 76% of the real cycling traffic variance (Raford et al., 2007). Tackling this challenge may reveal hotspots with subjective and objective risk parameters indicating particularly critical areas. From a cognitive perspective, the most interesting areas would be those that cyclists experience as dangerous (but that are not), as well as those where cyclists feel safe, but are frequently involved in accidents (for initial findings in this direction, see Schmidt and von Stülpnagel, 2018a,b). Given that such locations can be identified, it would be a worthwhile challenge to identify the sources of these discrepancies. Factors to look at include characteristics of the local infrastructure (e.g., traffic volume, different street sizes at intersections, or the availability of cycling lanes), but also a location's spatial configuration: An intersection might be constructed in such a way that the line of sight between a driver and a cyclists on a potential collision course is established too late (for an illustration, see "Collision Course: Why This Type Of Road Junction Will Keep Killing Cyclists, 2018). One way to investigate this matter are isovist analyses (e.g., Wiener et al., 2007; and see Schmidt and von Stülpnagel, 2018a,b; von Stülpnagel and Schmid, 2018, for an approach in this direction). Pinpointing crucial factors would contribute to our understanding of risk perception, but also bear the potential for new guidelines to design urban cycling structures both subjectively and objectively safer.

4.5. Conclusion

The increasing availability of non-academic VGI projects collecting information about perceived cycling risks promises new possibilities for the investigation and understanding of factors determining subjective risk during urban cycling. The present research provides evidence that parameters extracted from VGI as well as from authoritative data sources do represent valid indicators for the subjectively perceived risk of cycling at a specific location, even if their availability is opportunistic rather than scientifically guided. These data sources therefore provide valuable information for further investigations of subjective cycling risks.

Declaration of interests

None.

Acknowledgments

This research was supported by a seed grant of Freiburg University to the first author. We thank Nicolas Holland, Sonja Schmidt, and Laura Wächter for their support with data preparation and data collection. We thank Sueeddeutsche Zeitung and Mark-a-Spot for providing us access to the dataset investigated in Case 1. We thank Space Syntax Ltd for providing us with the axial map of Munich. We thank the people of "Besser unterwegs in Freiburg" for providing us access to the dataset investigated in Case 2. We thank the Freiburg police department for supporting us with cycling-related accident statistics. Data reported in this research is available upon request.

References

Abdalla, A., Weiser, P., 2011. Towards emotional and opinion-based layers in city GIS. Proceedings of Real Corp 2011.
Al-Sayed, K., Turner, A., Hillier, B., Iida, S., Penn, A., 2014. Space Syntax Methodology. Bartlett School of Architecture, UCL, London, UK.
Bafna, S., 2003. Space syntax: a brief introduction to its logic and analytical techniques. Environ. Behav. 35 (1), 17–29. https://doi.org/10.1177/0013916502238863.
Bishr, M., Kuhn, W., 2007. Geospatial information bottom-up: a matter of trust and semantics. Eur. Inf. Soc. 365–387.

Branion-Calles, M., Nelson, T., Winters, M., 2017. Comparing crowdsourced near-miss and collision cycling data and official bike safety reporting. Transp. Res. Rec.: J. Transp. Res. Board. 2662, 1–11. https://doi.org/10.3141/2662-01.

Brookhuis, K.A., de Waard, D., 2001. Assessment of drivers' workload: performance and subjective and physiological indexes. Stress, Workload and Fatigue.

Chataway, E.S., Kaplan, S., Nielsen, T.A.S., Prato, C.G., 2014. Safety perceptions and reported behavior related to cycling in mixed traffic: a comparison between Brisbane and Copenhagen. Transp. Res. Part F: Traffic Psychol. Behav. 23, 32–43. https://doi.org/10.1016/j.trf.2013.12.021.

Chaurand, N., Delhomme, P., 2013. Cyclists and drivers in road interactions: a comparison of perceived crash risk. Accid. Anal. Prev. 50, 1176–1184. https://doi.org/10.1016/j.aap.2012.09.005.

Cialdini, R.B., 2006. Influence: The Psychology of Persuasion (Rev. Ed., 1st Collins Business Essentials Ed). Harper Business, New York.

Collision Course: Why This Type Of Road Junction Will Keep Killing Cyclists, 2018. Retrieved May 3, 2018, from https://singletrackworld.com/2018/01/collision-course-why-this-type-of-road-junction-will-keep-killing-cyclists/.

Dubrovsky, V.J., Kiesler, S., Sethna, B.N., 1991. The equalization phenomenon: status effects in computer-mediated and face-to-face decision-making groups. Hum.—Comput. Interact. 6 (2), 119–146. https://doi.org/10.1207/s15327051hci0602_2.

Elwood, S., Goodchild, M.F., Sui, D.Z., 2012. Researching volunteered geographic information: spatial data, geographic research, and New social practice. Ann. Assoc. Am. Geogr. 102 (3), 571–590. https://doi.org/10.1080/00045608.2011.595657.

Encouraging Cycling in Central London | Space Syntax(2000). Retrieved December 5, 2014, from http://www.spacesyntax.com/project/encouraging-cycling-in-central-london/.

Fernández-Heredia, Á., Monzón, A., Jara-Díaz, S., 2014. Understanding cyclists' perceptions, keys for a successful bicycle promotion. Transp. Res. Part A: Policy Pract. 63, 1–11. https://doi.org/10.1016/j.tra.2014.02.013.

Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. GeoJournal 69 (4), 211–221. https://doi.org/10.1007/s10708-007-9111-y.

Goodchild, M.F., Li, L., 2012. Assuring the quality of volunteered geographic information. Spatial Stat. 1, 110–120. https://doi.org/10.1016/j.spasta.2012.03.002.

Gregoriades, A., Chrystodoulides, A., 2018. Extracting traffic safety knowledge from historical accident data. Adjunct Proceedings of the 14th International Conference on Location Based Services 109–114.

Haklay, M., 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey Datasets. Environ. Plann. B: Plann. Des. 37 (4), 682–703. https://doi.org/10.1068/b35097.

Hauthal, E., Burghardt, D., 2013. Extraction of location-based emotions from photo platforms. Progress in Location-Based Services. Springer, pp. 3–28.

Hillier, B., 1996. Space is the Machine: A Configurational Theory of Architecture. Cambridge University Press, Cambridge.

Hillier, B., Hanson, J., 1984. The Social Logic of Space. Cambridge University Press, Cambridge.

Hillier, B., Penn, A., Hanson, J., Grajewski, T., Xu, J., 1993. Natural movement: or, configuration and attraction in urban pedestrian movement. Environ. Plann. B: Plann. Des. 20 (1), 29–66. https://doi.org/10.1068/b200029.

Jahn, G., Oehme, A., Krems, J.F., Gelau, C., 2005. Peripheral detection as a workload measure in driving: effects of traffic complexity and route guidance system use in a driving study. Transp. Res. Part F: Traffic Psychol. Behav. 8 (3), 255–275. https://doi.org/10.1016/j.trf.2005.04.009.

Jessen, J., Jørgensen, A.H., 2011. Aggregated trustworthiness: redefining online credibility through social validation. First Monday 17 (1).

Jiang, B., 2015. Geospatial analysis requires a different way of thinking: the problem of spatial heterogeneity. GeoJournal 80 (1), 1–13. https://doi.org/10.1007/s10708-014-9537-y.

Leibovici, D.G., Rosser, J.F., Hodges, C., Evans, B., Jackson, M.J., Higgins, C.I., 2017. On data quality assurance and conflation entanglement in crowdsourcing for environmental studies. ISPRS Int. J. Geo-Inf. 6 (3), 78. https://doi.org/10.3390/ijgi6030078.

Meek, S., Jackson, M., Leibovici, D., 2014. A flexible framework for assessing the quality of crowdsourced data. Presented at the AGILE'2014.

Molino, J.A., Kennedy, J.F., Johnson, P.L., Beuse, P.A., Emo, A., Do, A., 2009. Pedestrian and Bicyclist Exposure to Risk: Methodology for Estimation in an Urban Environment (With Discussion and Closure).

Nelson, T.A., Denouden, T., Jestico, B., Laberee, K., Winters, M., 2015. BikeMaps.Org: a global tool for collision and near miss mapping. Front. Public Health 3. https://doi.org/10.3389/fpubh.2015.00053.

Parkin, J., Wardman, M., Page, M., 2007. Models of perceived cycling risk and route acceptability. Accid. Anal. Prev. 39 (2), 364–371. https://doi.org/10.1016/j.aap.2006.08.007.

Raford, N., Chiaradia, A., Gil, J., 2007. Space syntax: the role of urban form in cyclist route choice in central London. Safe Transp. Res. Educ. Cent.

Rietveld, P., Daniel, V., 2004. Determinants of bicycle use: do municipal policies matter? Transp. Res. Part A: Policy Pract. 38 (7), 531–550. https://doi.org/10.1016/j.tra.2004.05.003.

Sanders, R.L., 2015. Perceived traffic risk for cyclists: the impact of near miss and collision experience. Accid. Anal. Prev. 75, 26–34. https://doi.org/10.1016/j.aap.2014.11.004.

Schepers, P., Hagenzieker, M., Methorst, R., van Wee, B., Wegman, F., 2014. A conceptual framework for road safety and mobility applied to cycling safety. Accid. Anal. Prev. 62 (Suppl. C), 331–340. https://doi.org/10.1016/j.aap.2013.03.032.

Schmid, K., von Stülpnagel, R., 2018a. Quantifying local spatial properties through LiDAR-based isovists for an evaluation of opinion-based VGI in a VR setup. Adjunct Proceedings of the 14th International Conference on Location Based Services 173–178.

Schmidt, S., von Stülpnagel, R., 2018b. Risk perception and gaze behavior during urban cycling—a field study. Eye Tracking for Spatial Research, Proceedings of the 3rd International Workshop.

Schwering, A., 2008. Approaches to semantic similarity measurement for geo-spatial data: a survey. Trans. GIS 12 (1), 5–29. https://doi.org/10.1111/j.1467-9671.2008.01084.x.

Sørensen, M., Mosslemi, M., 2009. subjective and objective safety. The effect of road safety measures on subjective safety among vulnerable road users. TOI Rep. 1009.

Spielman, S.E., 2014. Spatial collective intelligence? Credibility, accuracy, and volunteered geographic information. Cartogr. Geogr. Inf. Sci. 41 (2), 115–124. https://doi.org/10.1080/15230406.2013.874200.

von Stülpnagel, R., Schmid, K., submitted for publication. Risk perception during urban cycling in volunteered geographic information and in the lab: effects of the vista space's spatial properties.

Wardman, M., Hatfield, R., Page, M., 1997. The UK national cycling strategy: can improved facilities meet the targets? Transp. Policy 4 (2), 123–133.

Washington, S., Haworth, N., Schramm, A., 2012. Relationships between self-reported bicycling injuries and perceived risk of cyclists in Queensland, Australia. Transp. Res. Rec.: J. Transp. Res. Board. 2314, 57–65. https://doi.org/10.3141/2314-08.

Wiener, J.M., Franz, G., Rossmanith, N., Reichelt, A., Mallot, H.A., Bülthoff, H.H., 2007. Isovist analysis captures properties of space relevant for locomotion and experience. Perception 36 (7), 1066–1083. https://doi.org/10.1068/p5587.