Data in Brief 54 (2024) 110278

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Using demographic, psychosocial, behavioural and safety-related factors to assess cyclists' behaviour: A comparative approach across 19 countries

Sergio A. Useche[a,*], Francisco Alonso[a], Aleksey Boyko[b], Polina Buyvol[b], Isaac Castañeda[c], Boris Cendales[d], Arturo Cervantes[c], Tomas Echiburu[e], Mireia Faus[a], Zuleide Feitosa[f], Cesareo Fernandez[g], Javier Gene-Morales[a], Bas de Geus[h], Jozef Gnap[i], Mohd K. Ibrahim[j], Kira H. Janstrup[k], Ignacio Lijarcio[a], Irina Makarova[b], Miroslava Mikusova[i], Mette Møller[k], Sylvain Ngueuteu-Fouaka[l], Steve O'Hern[m], Mauricio Orozco-Fontalvo[n], German Rojas[d], Ksenia Shubenkova[b], Felix Siebert[k], Jose Soto[e], Amanda N. Stephens[o], Yonggang Wang[p], Elias Willberg[q], Phillip Wintersberger[r], Yongzheng Yang[s], Linus Zeuwts[t], Zarir H. Zulkipli[h], Rich McIlroy[u]

[a] *University of Valencia, Valencia, Spain*
[b] *Kazan Federal University, Kazan, Russia*
[c] *Anahuac University, Mexico*
[d] *El Bosque University, Bogotá, Colombia*
[e] *Universidad Católica de Chile, Santiago, Chile*
[f] *Universidade de Brasilia, Brasilia, Brazil*
[g] *Jaume I University, Castellon, Spain*
[h] *University of Louvain, Belgium*
[i] *University of Žilina, Bratislava, Slovakia*
[j] *Malaysian Institute of Road Safety Research, Kajang, Malaysia*
[k] *Technical University of Denmark, Copenhague, Denmark*
[l] *Université de Dschang, Dschang, Cameroon*
[m] *University of Leeds, United Kingdom*
[n] *Universidade de Lisboa, Lisboa, Portugal*
[o] *Monash University, Melbourne, Australia*
[p] *Chang'an University, Chang'an, China*
[q] *University of Helsinki, Helsinki, Finland*

[*] Corresponding author.
E-mail address: sergio.useche@uv.es (S.A. Useche).

https://doi.org/10.1016/j.dib.2024.110278

[r] Technical University of Wien, Wien, Austria
[s] Wuhan University, China
[t] Ghent University, Ghent, Belgium
[u] University of Southampton, Southampton, England

ARTICLE INFO

Article history:
Received 10 November 2023
Revised 23 January 2024
Accepted 26 February 2024
Available online 4 March 2024

Keywords:
Cycling
Riding behavior
CBQ
Cycling safety
Multi-country research

ABSTRACT

This Data in Brief (DiB) article presents the differences in cycling behaviors related to violations, errors, and positive behaviors by region. The study data were collected by means of a structured questionnaire applied to a full sample of 7,001 participants from 19 countries, distributed over 5 continents. This paper proposes descriptive statistics, as well as common statistical tests. The aim is to enable authors to make their own analyses, not to provide precise interpretations. For further information about the macro project supporting the collection of these data, it is advised to refer to the paper titled "Cross-culturally approaching the cycling behavior questionnaire (CBQ): Evidence from 19 countries", published in *Transportation Research Part F: Traffic Psychology and Behavior*.

Specifications Table

Subject	Psychology
Specific subject area	Road safety; Mobility; Cycling behavior.
Data format	Filtered and Analyzed
Type of data	- Tables with the data of the differences of the CBQ questionnaire factors according to region. - Graphs representing the Cycling Behavior Questionnaire (CBQ) factors scored and a Pearson correlation analysis of the study variables. - Database.
Data collection	Data collection was performed through an electronic survey, using an online questionnaire translated to each country's most spoken language. There was used a convenience (pseudo-probabilistic) sampling method, aimed at gathering data from current (and active) bicycle users. Regarding the recruitment tactics, various initiatives were carried out, such as promotions on social media, distribution of questionnaires in classrooms, use of mailing lists and collaboration with national cycling federations, with the aim of collection information in all countries.
Data source location	Data were collected in 19 countries: Australia, Austria, Belgium, Brazil, Cameroon, Chile, China, Colombia, Denmark, Dominican Republic, Finland, Germany, Malaysia, Mexico, Poland, Russia, Slovakia, Spain and United Kingdom.
Data accessibility	The appended data are directly derived from the original database reported in the article. It contains the .CSV format dataset, the root questionnaire, and the study codebook. Data are fully accessible at the permanent link (Dataverse): https://doi.org/10.7910/DVN/EP6QLN These data can be freely used for research, divulgation, and scientific purposes, as long as the original source is acknowledged.
Related research article	S.A. Useche, F. Alonso, A. Boyko, P. Buyvol, I. Castañeda, B. Cendales, B., A. Cervantes, T. Echiburu, M. Faus, Z. Feitosa, J. Gene, A. Gonzalez-Marin, V. Gonzalez, J. Gnap, M.K. Ibrahim, K.H. Janstrup, A. Javadinejad, I. Makarova, R. McIlroy, M. Mikusova, M. Møller, S. Ngueuteu-Fouaka, S. O'Hern, M. Orozco-Fontalvo, K. Shubenkova, F. Siebert, J. Soto, A.N. Stephens, R. Valle-Escolano, Y. Wang, E. Willberg, P. Wintersberger, L. Zeuwts, Z.H. Zulkipli, L. Montoro. Cross-culturally approaching the cycling behaviour questionnaire (CBQ): Evidence from 19 countries. Transp Res Part F Traffic Psychol Behav. (2022) 91: 386–400. doi: 10.1016/j.trf.2022.10.025 [1]

1. Value of the Data

- This article provides information about the differences in the behavior of cycle users according to their region of residence.
- The appended dataset provides information on the Cycling Behavior Questionnaire (CBQ) instrument, making it possible to analyze its outcomes in different regions of the world.
- These data could be used by several stakeholders, including researchers and policymakers in the countries involved, for designungy plans and strengthening safety interventions covering urban cyclists.
- As for practical usefulness, the information contained in the dataset could be useful for designing awareness campaigns on common risk behaviors in the cycling population and encouraging positive behaviors in this vulnerable group of road users.
- This database can also benefit the work of teachers or lecturers in statistics and psychology courses for student instruction, especially in the case of those specialists addressing issues such as user behavior, active transport, and safe mobility.

2. Objective

The aim of this study was to analyze the differences in cycling behavior (including violations, errors and positive behaviors) according to their region of residence (Europe, Oceania, Latin America, Asia and Africa), using the Cycling Behavior Questionnaire (CBQ) instrument applied in 19 countries.

3. Data Description

The dataset in this article provides information on differences in cycling behavior as a function of the region of residence, especially in relation to traffic violations, errors, and positive riding behaviors. There is a complete overview of the outcomes of a large application of the Cycling Behavior Questionnaire (CBQ) in its 29-item version [1,2], alongside the Risk Perception and Regulation Scale [1,3,4] and supplementary items addressing self-reported crash-related cycling incidents suffered in a time span of five (5) years, which have been categorized both in a continuous form (i.e., counting the total number of S-R crashes) and dichotomized (i.e., yes/no) for logit and categorical analyses. Both individual item and subscale scores have been calculated, standardized, and labeled for this database.

In addition to the .CSV file, this data article contains two online appendixes available here: a root questionnaire form (Appendix I) and its codebook (Appendix II), added to give further researchers the possibility of comprehensively labeling and interpreting the study variables composing the dataset file.

Table 1 shows the ANOVA results and Fig. 1 presents the standardized scores obtained for the three CBQ factors in the 5 regions evaluated. Tables 2–4 display the differences in cycling behavior as a function of region, each table corresponding to each of the factors analyzed through the Tukey post hoc test. Finally, Fig. 2 shows a bivariate Pearson correlation matrix between the study variables, including the relationships between cyclists' violations, errors, and positive behaviors with knowledge of regulations, risk perception, and self-reported accidents (i.e., cycling safety-related incidents) as bicycle users.

The inclusion of these descriptive statistical outcomes serves the purpose of providing the reader with a preliminary overview of the potential influence of demographic variables on cycling behavioral outcomes, as well as their bivariate associations. This information might facilitate subsequent analyses to be performed with this behavioral database, such as age, gender, and country-based comparisons, or predictive models aiming at explaining their effect on safety outcomes.

Table 1
Mean comparisons of the CBQ factors in different regions.

CBQ Factor	Test setting	Table sum	gl	Quadratic mean	F	Sig.
Traffic Violations	Between Groups	439.74	4	109.93	117.23	<0.001
	Within Groups	6560.25	6996	.938		
	Total	7000.00	7000			
Errors	Between Groups	803.10	4	200.77	226.66	<0.001
	Within Groups	6196.89	6996	.886		
	Total	7000.00	7000			
Positive Behaviors	Between Groups	141.31	4	35.32	36.03	<0.001
	Within Groups	6858.68	6996	.980		
	Total	7000.00	7000			

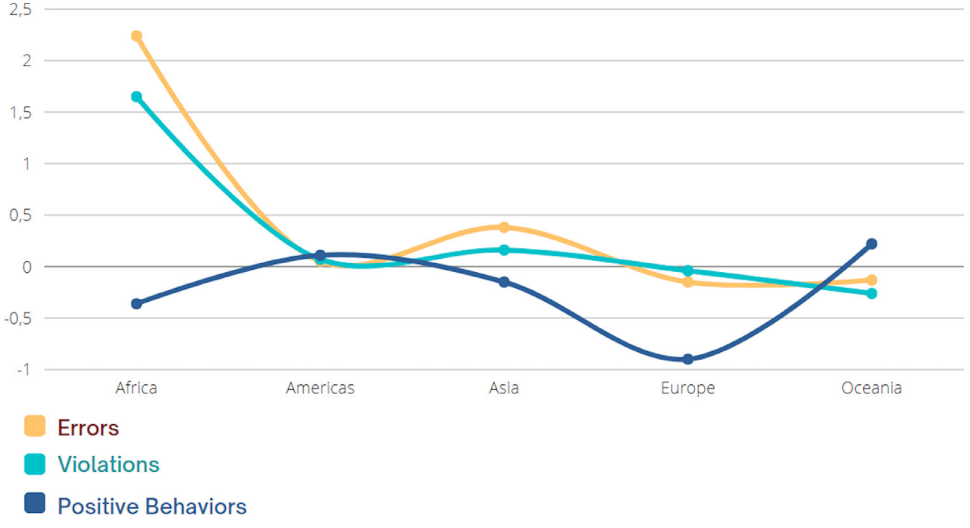

Fig. 1. Comparative scores in the CBQ factors according to cyclists' region of provenance. Note: The variable scores have been standardized in order to favor their fair comparability.

4. Experimental Design, Materials and Methods

4.1. Participants

In this cross-sectional study, the sample consisted of a total of 7001 participants from 19 countries in 5 regions, which were: Australia, Belgium, Brazil, Cameroon, Chile, China, Colombia, Denmark, Dominican Republic, Finland, Germany, Malaysia, Mexico, Poland, Russia, Slovakia, Spain, and the United Kingdom. The average age of the study participants was 36.63 years (SD = 14.17) [range 16–83], and the gender ratio was 38.5% females and 60.8% males.

Data were gathered through an electronic survey translated into the most spoken language of each country, from which data from at least one houndred of current and active bicycle users per country were gathered. Even though the e-survey utilized a consistently pre-structured questionnaire format based on Google Forms, the application platform employed for data collection slightly varied in some few countries. This discrepancy primarily arose due to convenience or institutional recommendations favoring specific paid survey platforms (e.g., SurveyXact and Qualtrics in Danish and Australian universities) or limitations imposed by country-specific unavailability and data restrictions (e.g., Google Forms being a prohibited website in China).

Table 2
Post-Hoc (Tukey HSD) analysis – mean comparisons for CBQ Violations. Factor: Cyclists' region of provenance.

Region (I)	Region (J)	Mean Diff. (I-J)	Std. Error	Sig (p)	95% CI	
					Lower	Upper
Europe	Latin America	−0.12*	.02	<0.001	−0.19	−0.04
	Asia	−0.20*	.04	<0.001	−0.31	−0.10
	Oceania	.21*	.03	<0.001	.12	.30
	Africa	−1.70*	.09	<0.001	−1.95	−1.45
Latin America	Europe	.12*	.03	<0.001	.04	.19
	Asia	−0.09	.04	.23	−0.20	.02
	Oceania	.33*	.03	<0.001	.23	.43
	Africa	−1.58*	.09	<0.001	−1.83	−1.33
Asia	Europe	.21*	.03	<0.001	.10	.31
	Latin America	.08	.04	.23	−0.02	.20
	Oceania	.42*	.04	<0.001	.29	.55
	Africa	−1.49*	.09	<0.001	−1.75	−1.23
Oceania	Europe	−0.21*	.03	<0.001	−0.30	−0.12
	Latin America	−0.33*	.03	<0.001	−0.43	−0.23
	Asia	−0.42*	.04	<0.001	−0.55	−0.29
	Africa	−1.91*	.09	<0.001	−2.17	−1.65
Africa	Europe	1.70*	.09	<0.001	1.45	1.95
	Latin America	1.58*	.09	<0.001	1.33	1.83
	Asia	1.49*	.09	<0.001	1.23	1.75
	Oceania	1.92*	.09	<0.001	1.66	2.17

Table 3
Post-Hoc (Tukey HSD) analysis – mean comparisons for CBQ Errors. Factor: Region of provenance.

Region (I)	Region (J)	Mean Diff. (I-J)	Std. Error	Sig (p)	95% CI	
					Lower	Upper
Europe	Latin America	−0.20*	.02	<0.001	−0.27	−0.12
	Asia	−0.53*	.03	<0.001	−0.64	−0.42
	Oceania	−0.02	.03	.971	−0.10	.06
	Africa	−2.39*	.08	<0.001	−2.63	−2.15
Latin America	Europe	.20*	.02	<0.001	.12	.27
	Asia	−0.33*	.04	<0.001	−0.44	−0.21
	Oceania	.18*	.03	<0.001	.08	.28
	Africa	−2.19*	.08	<0.001	−2.43	−1.94
Asia	Europe	.53*	.03	<0.001	.42	.64
	Latin America	.33*	.04	<0.001	.21	.44
	Oceania	.51*	.04	<0.001	.39	.63
	Africa	−1.85*	.09	<0.001	−2.11	−1.60
Oceania	Europe	.020	.03	.971	−0.06	.10
	Latin America	−0.18*	.03	<0.001	−0.28	−0.08
	Asia	−0.51*	.04	<0.001	−0.63	−0.39
	Africa	−2.37*	.09	<0.001	−2.62	−2.12
Africa	Europe	2.39*	.08	<0.001	2.15	2.63
	Latin America	2.19*	.08	<0.001	1.94	2.43
	Asia	1.85*	.09	<0.001	1.60	2.11
	Oceania	2.37*	.09	<0.001	2.12	2.62

In terms of recruitment strategies, various enrollment methods were implemented to collect data across all countries. This included social media advertising, sharing questionnaires in classrooms, utilizing mailing lists and participants from previous research endeavors (utilizing an 'invitation+reminder' approach), and collaborating with national cyclist federations. It's worth noting that participants in the study were not provided with any financial incentives, and no commercial panels were used. Although the sample was very large, it is essential to remark that these particularities should be considered as potential sources of bias when interpreting the data. The nature and dynamics of these recruitment methods may inadvertently favor specific

Table 4
Post-Hoc (Tukey HSD) analysis – mean comparisons for CBQ Positive Behavioirs. Factor: Region of provenance.

Region (I)	Region (J)	Mean Diff. (I-J)	Std. Error	Sig (p)	95% CI	
					Lower	Upper
Europe	Latin America	−0.20*	.02	<0.001	−0.28	−0.12
	Asia	.066	.04	.470	−0.04	.17
	Oceania	−0.31*	.03	<0.001	−0.41	−0.22
	Africa	.26*	.09	.033	.01	.52
Latin America	Europe	.20*	.02	<0.001	.12	.28
	Asia	.27*	.04	<0.001	.15	.38
	Oceania	−0.11*	.03	.010	−0.21	−0.01
	Africa	.47*	.09	<0.001	.21	.72
Asia	Europe	−0.06	.04	.470	−0.17	.04
	Latin America	−0.27*	.04	<0.001	−0.38	−0.15
	Oceania	−0.38*	.04	<0.001	−0.51	−0.25
	Africa	.20	.09	.231	−0.06	.46
Oceania	Europe	.31*	.03	<0.001	.22	.41
	Latin America	.11*	.03	.010	.01	.21
	Asia	.38*	.04	<0.001	.25	.51
	Africa	.58*	.09	<0.001	.32	.84
Africa	Europe	−0.26*	.09	.033	−0.52	−0.01
	Latin America	−0.47*	.09	<0.001	−0.72	−0.21
	Asia	−0.20	.09	.231	−0.46	.06
	Oceania	−0.58*	.09	<0.001	−0.84	−0.32

segments of the cycling population, introducing a potential confounding source. This is worth considering when drawing conclusions and rationales based on the study outcomes.

The overall response rate (i.e., questionnaires completed once the study link was accessed) of this study was estimated at around 46%, even accounting for approximately 265 system-reported incomplete forms (fulfillment rate < 80%), and about 20 duplicate MAC addresses, which were subsequently discarded from the final dataset.

4.2. Questionnaire

The study data were collected through an electronic (online) questionnaire, previously translated and re-translated 'to and from' the most widely spoken language(s) in each country by qualified researchers, appending the scales described below:

The questionnaire included the Cycling Behavior Questionnaire (CBQ) [1], which addresses driving behaviors through a three-dimensional structure composed of 29 items distributed in three sub-scales: Traffic violations (Factor 1; 8 items), Errors (Factor 2; 15 items), and Positive behaviors (Factor 3; 4 items). In operational terms, factors 1 and 2 correspond to hazardous riding expressions. However, while factor 1 (traffic violations) refers to deliberate risky behaviors (e.g., running a red light), the items of factor 2 (cycling errors) make reference to unintentional behaviors that regularly increase cyclists' safety risks (e.g., failing to notice the presence of pedestrians while turning). On the other hand, factor 3 (positive behaviors) refers to safety-increasing habits that decrease their involvement in risk-related scenarios (e.g., avoiding cycling under adverse weather conditions) [5].

Supplementarily, and with the aim of facilitating variable crossings and validity analyses, there was included the Risk Perception and Regulation Scale (RPRS) [3], a 12-item self-report questionnaire composed of two factors: risk perception (Factor 1; 7 items) assessing the perceived risk regarding typical road safety issues (e.g., infrastructure flaws) and traffic rule knowledge (Factor 2; 5 items) prompting cyclists to self-report their awareness of basic traffic norms and conventions (e.g., traffic signal awareness). Finally, a safety-related indicator was incorporated, asking participants to indicate the number of accidents or crash-related incidents in which they have been involved as cyclists, regardless of their severity.

Fig. 2. Graphical bivariate correlations between factors included in the dataset.

Limitations

Although this study used validated research tools and gathered a large sample of cyclists, there are some key limitations to acknowledge and discuss, given that they can biase the study outcomes and/or their subsequent interpretations.

Participant recruitment processes varied between countries, and there was a lack of representation in some regions, particularly in Africa. This can be attributed to the limited presence of research network representatives and connected technologies, a small number of regular cyclists, and their relatively lesser 'maturity' in both cycling and applied research dynamics in that region. In addition, the sampling strategy could not be uniform across all countries, and data were collected following a pseudo-probabilistic convenience sample method, aimed at gathering data in strategic collectives (i.e., active urban cycling populations) of these countries.

The observed gender disparity within the sample population, with a female-to-male cyclist ratio of 1:1.57, brings attention to two significant issues. Firstly, despite a growing trend in female involvement in urban cycling, the numbers still fall short of achieving parity with male riders. Secondly, contemporary gender concerns come to the forefront, as the study includes a substantial number of male ($n = 4255$) and female participants ($n = 2696$), meeting the mini-

mum size requirements for conducting weighted comparative analyses under optimal statistical assumptions. However, non-binary participants ($n = 50$), apart from underrepresented in this study, are overall scarcely addressed in current active transport studies [6]. One potential solution suggested by recent literature to address this issue might be the use of intentional or stratified sampling procedures explicitly pursuing to cover underrepresented segments of population, something useful to increase the participation of groups that may be statistically less conventional in terms of cycling representation and visibility [7].

Ethics Statement

The contribution of this study's partakers was voluntary and anonymous. After the explanation of the research objectives and considerations related to the research, all participants gave their informed consent before participating in the study (a copy of the informed consent is available alongside this paper). Underaged participants were requested to obtain informed consent from their parents or tutors ("*on behalf of*"). The study was approved by the Ethics Committee of the Traffic and Road Safety Research Institute of the University of Valencia (IRB approval number: HE0001241120).

Credit Author Statement

Sergio A. Useche: Visualization, Conceptualization, Supervision, Data curation, Investigation, Data analysis; Writing – Original draft; Writing- Reviewing and Editing; **Francisco Alonso:** Data curation, Investigation, Resources; **Aleksey Boyko:** Data curation, Investigation; **Polina Buyvol:** Data curation, Investigation; **Isaac Castañeda:** Data curation, Investigation; **Boris Cendales:** Data curation, Investigation; **Arturo Cervantes:** Data curation, Investigation; **Tomas Echiburu:** Data curation, Investigation; **Mireia Faus:** Data curation, Investigation; **Zuleide Feitosa:** Data curation; **Cesareo Fernandez:** Data curation, Investigation; **Javier Gene-Morales:** Data curation, Investigation, Writing- Reviewing; **Bas de Geus:** Investigation; **Jozef Gnap:** Data curation, Investigation; **Mohd K. Ibrahim:** Data curation, Investigation, Writing- Reviewing; **Kira H. Janstrup:** Data curation, Investigation; **Ignacio Lijarcio:** Data curation, Investigation; **Irijna Makarova:** Data curation, Investigation; **Miroslava Mikusova:** Data curation, Investigation; **Mette Møller:** Data curation, Investigation; **Sylvain Ngueuteu-Fouaka:** Data curation, Investigation; **Steve O'Hern:** Data curation, Investigation; **Mauricio Orozco-Fontalvo:** Data curation, Investigation; **German Rojas:** Data curation, Investigation; **Ksenia Shubenkova:** Data curation, Investigation; **Felix Siebert:** Data curation, Investigation; **Jose Soto:** Data curation, Investigation; **Amanda N. Stephens:** Data curation, Investigation; **Yonggang Wang:** Data curation, Investigation; **Ellias Willberg:** Data curation, Investigation; **Phillip Wintersberger:** Data curation, Investigation; **Yongzheng Yang:** Investigation; **Linus Zeuwts:** Data curation, Investigation; **Zarir H. Zulkipli:** Data curation, Investigation; **Rich McIlroy:** Data curation, Investigation, Writing- Reviewing.

Data availability

Raw Study Data (Original data) (Dataverse)

Acknowledgements

The authors wish to thank all participants of the study and Lucía Giménez-Tornero (certified translator) for the professional edition of the final version of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2024.110278.

References

[1] S.A. Useche, F. Alonso, A. Boyko, P. Buyvol, I. Castañeda, B. Cendales, B., A. Cervantes, T. Echiburu, M. Faus, Z. Feitosa, J. Gene, A. Gonzalez-Marin, V. Gonzalez, J. Gnap, M.K. Ibrahim, K.H. Janstrup, A. Javadinejad, I. Makarova, R. McIlroy, M. Mikusova, M. Møller, S. Ngueuteu-Fouaka, S. O'Hern, M. Orozco-Fontalvo, K. Shubenkova, F. Siebert, J. Soto, A.N. Stephens, R. Valle-Escolano, Y. Wang, E. Willberg, P. Wintersberger, L. Zeuwts, Z.H. Zulkipli, L. Montoro, Cross-culturally approaching the cycling behaviour questionnaire (CBQ): evidence from 19 countries, Transp. Res. Part F 91 (2022) 386–400, doi:10.1016/j.trf.2022.10.025.
[2] S.A. Useche, L. Montoro, J.M. Tomas, B. Cendales, Validation of the Cycling Behavior Questionnaire: a tool for measuring cyclists' road behaviors, Transp. Res. Part F 58 (2018) 1021–1030, doi:10.1016/j.trf.2018.08.003.
[3] S.A. Useche, L. Montoro, F. Alonso, F.M. Tortosa, Does gender really matter? A structural equation model to explain risky and positive cycling behaviors, Accid. Anal. Prev. 118 (2018) 86–95, doi:10.1016/j.aap.2018.05.022.
[4] R. McIlroy, S.A. Useche, A. Gonzalez-Marin, To what extent do our walking and cycling behaviours relate to each other, and do we cycle as well as we think we do? Piloting the Walking and Cycling Behaviour Questionnaires in the UK, Accid. Anal. Prev. 168 (2022) 106597, doi:10.1016/j.aap.2022.106597.
[5] M. Møller, S.A. Useche, F. Siebert, K.H. Janstrup, What differentiates e-bike riders from conventional cyclists in Denmark? A user-based study, J. Transp. Health 34 (2024) 101748, doi:10.1016/j.jth.2023.101748.
[6] B. Braumüller, T. Menzel, I. Hartmann-Tews, Gender identities in organized sports-athletes' experiences and organizational strategies of inclusion, Front. Sociol. 5 (2020) 578213, doi:10.3389/fsoc.2020.578213.
[7] P. Christidis, E. Navajas Cawood, D. Fiorello, Challenges for urban transport policy after the Covid-19 pandemic: main findings from a survey in 20 European cities, Transp. Policy 129 (2022) 105–116, doi:10.1016/j.tranpol.2022.10.007.