A Naturalistic Approach for Identifying Latent Behavioral Classes of Two- And Three-Wheel Vehicle Users

Cailis Bullard[*1], Sonja Haustein[1], Ravi Seshadri[1], Oscar Oviedo-Trespalacios[2] & Felix Wilhelm Siebert[1]

[1]Department of Technology, Management, and Economics, Technical University of Denmark, Lyngby, Denmark

[2]Department of Technology, Policy, and Management, Delft University of Technology, Delft, Netherlands

SHORT SUMMARY

Two- and three-wheel vehicle users are generally considered a homogeneous group in traffic safety research. However, evidence supports that rider behaviors vary largely based on situational factors such as interactions with vehicles, other riders, and the environment. Identifying latent behavioral subgroups can therefore provide a more meaningful basis for understanding behavioral patterns among riders, as well as developing targeted safety interventions. This study aims to uncover behavioral patterns among two- and three-wheel vehicle users within Copenhagen's bicycle lanes. Observational video data collected in 2025 at four locations were analyzed using Latent Class Analysis to identify distinct behavioral groups based on vehicle, environmental, and sociodemographic factors. Three resulting classes were then described based on the profile of their observed characteristics. Findings from this research provide insights that can be used to inform targeted safety and policy interventions as a means of addressing subgroup specific behaviors among two- and three-wheel vehicle users.

Keywords: Cyclist behavior, Distracted riding, Latent class analysis, Naturalistic observational data, Traffic Safety

1. INTRODUCTION

Within traffic safety research, two- and three-wheel vehicle users are frequently treated as though they are a homogeneous group exhibiting similar behaviors across most situations on the road and are analyzed as such (Schepers et al., 2014). However, a growing body of literature suggests these riders are not a homogeneous group, emphasizing that even within individual riders, behavior varies strongly depending on situation specific factors such as rider-to-vehicle interactions, rider-to-rider interactions, and their environment (de Freitas et al., 2025; Aydar et al., 2025; Griswold et al., 2018). Additionally, many safety related behaviors such as phone distraction, varied levels of vehicle control, and interaction patterns between riders and their environment are analyzed in relation only to individual rider characteristics (Useche et al., 2018). However, addressing such behaviors at the individual level is impractical and would likely lead to little improvement in overall safety. Indeed, research shows that there are often clusters within a road user group that have similar characteristics on multiple variables and subsequently are found to also show similar behaviors (Nguyen-Phuoc et al., 2020), which would be more feasible to address in practice. The lack of identification of latent behavioral groupings is a strong barrier for understanding differences in risk exposure among riders and thus limits the development of targeted safety interventions based on the sociodemographic and behavioral patterns within said groups. For example, one group of riders may frequently use their phone and ride with one hand, while another mainly rides in dense, platooned traffic but maintains control. These situations involve different types of risk, and identifying such groups helps make this visible and supports more targeted safety interventions (Fisa et al., 2022).

Self-reported approaches have been widely used to identify rider behavior subgroups and provide valuable insights into attitudes, perceptions, and motivations (Ramos-Ruiz et al., 2025; Damant-Sirois et al., 2014; Dill & McNeil, 2013). However, like all methods, they have limitations, particularly in capturing moment-to-moment, context-dependent behaviors and naturalistic interactions (Eby, 2011). Furthermore, survey data are prone to underreporting in both distracted and safety critical behaviors due to reporting bias (Friedman, 1993; Useche et al., 2018) and social desirability bias (Grimm, 2010; O'Hern et al., 2026). As such, there is a need to incorporate complementary approaches, such as observational methods, to capture how behaviors unfold in real-world, context-dependent settings. In this research, we adopt an observational approach to capture real-world, context-dependent rider behaviors such as phone use, steering control, and interactions with traffic that are difficult to reliably assess through self-reported methods.

The aim of this study is primarily to uncover the behavioral patterns of two- and three-wheel vehicle users present in the Copenhagen bicycle lanes. To this end, observational video data that was collected in 2025 across four separate locations in Copenhagen before a Latent class analysis (LCA) approach is employed to produce distinct behavioral classes of individuals from said classes. These classes are then subsequently described using the profiles of the indicators and covariates utilized in the final selected model. Such findings present an opportunity to inform both policy makers and target road safety campaigns toward specific groups of the riders, improving overall safety and awareness of riders in the bicycle lanes. Furthermore, findings from this research are easily extendable to other regions wherein there exist similar modal shares of these road users and geometrically similar bicycle lanes.

2. METHODOLOGY

This section describes the data collection, as well as the analytical approach taken to analyze the data for the purposes of this study.

Data collection

Naturalistic observational video data were collected across four different separated bicycle lane locations within Copenhagen. Data was collected between May and June of 2025, in consultation with the Copenhagen Municipality's Road Authority. Observation sites were selected based on their similar geometric features and the level of potential for pedestrian interaction present at said locations, i.e. if there was an adjacent pedestrian path to the bicycle lane. Sites were classified as: "non-interaction" (Site 1 and Site 2) where there is no adjacent pedestrian path to the bicycle lane, and "interaction" (Site 3 and Site 4) where there is an adjacent pedestrian path to the bicycle lane (Figure 1). This measure was taken to observe variations in two- and three-wheel vehicle user behavior at locations with varying traffic environments.

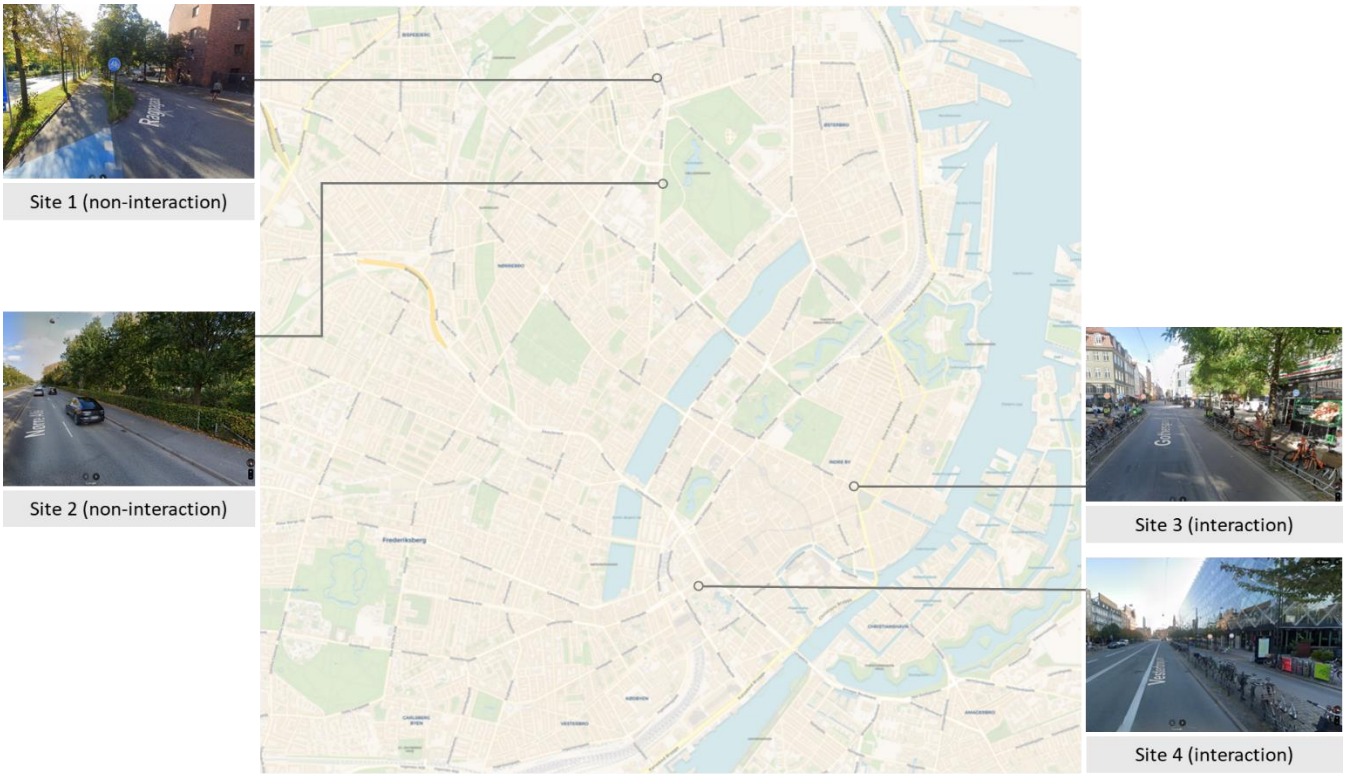


Figure 1. Observational site locations in Copenhagen, Denmark

Video data was collected using a two-camera setup where both were mounted at the edges of the bicycle lane. The heights and angles used in this observational study design were chosen to capture a clear view of the individuals' behaviors and overall traffic conditions/behaviors and are based on the experience of prior studies on bicycle lane behavior (Siebert et al., 2021). Vehicle types examined in this study include bicycles, mopeds, e-bikes, e-scooters, cargo bikes, and others. Data collection was carried out for one day at each of the four locations, during weekdays and within daylight hours (5:00 to 21:00). To reduce manual behavior coding to a manageable level, videos were annotated for the first 15 minutes of every 2-hour interval beginning at 05:00 and concluding at 19:00, which amounts to approximately 18% of the total amount of video data collected across all sites. The chosen approach to video annotation was expected to provide an overview of temporal trends in both vehicle user and traffic behavior at each site. Manual annotations resulted in exactly 2,367 observations of individual two- and three-wheel road users in the bicycle lane.

After the processing and cleaning of data, 2,167 observations were left for analysis. The observed characteristics considered in this study are presented in Table 1.

Table 1. Characteristics of Two- and Three-Wheel Vehicle User Observations

Variable		Frequency	Proportion
Total observations		2167	
Steering control	Two hands steering	1948	89.9%
	One hand steering	210	9.7%
	No hands steering	9	0.4%
Obstruction of controls (e.g. bag, etc.)	No	2122	97.9%
	Yes	46	2.1%
Headphone use	No	1481	68.3%
	Yes	667	30.8%
	Unknown	19	0.9%
Phone use	No	2109	97.3%
	Mounted phone use	17	0.8%
	Phone in hand	41	1.9%
Normal Platooning	No	1094	50.5%
	Yes	1069	49.3%
	Unknown	4	0.2%
Social Platooning	No	2113	97.5%
	Yes	54	2.5%
Traffic in bike lane	Single vehicles in bike lane	685	31.6%
	≤2 vehicles in bike lane	897	41.4%
	≤4 vehicles in bike lane	324	15.0%
	>4 vehicles in bike lane	261	12.0%
Helmet use	No helmet use	1421	65.6%
	Helmet present, in use	737	34.0%
	Helmet present, not in use	9	0.4%
Brake configuration	Right- and left-hand brake	901	41.6%
	Right-hand brake only	213	9.8%
	Left-hand brake only	1053	48.6%
Rider age	1-12 yrs	29	1.3%
	13-18 yrs	90	4.2%
	19-24 yrs	392	18.1%
	25-44 yrs	1219	56.3%
	45-64 yrs	356	16.4%
	>65 yrs	81	3.7%
Rider gender	Male	1131	52.2%
	Female	1036	47.8%
Handlebar configuration	Flat bar	744	34.3%

	Flat raise bar	622	28.7%
	Drop bar	147	6.8%
	Cruiser	649	29.9%
	Bull horn bar	5	0.2%
Time of day	05:00-0515	13	0.6%
	07:00-07:15	339	15.6%
	09:00-09:15	372	17.2%
	11:00-11:15	180	8.3%
	13:00-13:15	225	10.4%
	15:00-15:15	372	17.2%
	17:00-17:15	470	21.7%
	19:00-19:15	196	9.0%
Vehicle design	City bike	1860	85.8%
	Road bike	140	6.5%
	Mountain bike	44	2.0%
	Cargo bicycle	113	5.2%
	Scooter	7	0.3%
	Motorcycle (Moped)	4	0.2%
Vehicle function	Personal	1978	91.3%
	Extended rental	58	2.7%
	Shared	126	5.8%
	Taxi	3	0.1%
	Delivery	1	0.0%
Vehicle power	None	1879	86.7%
	Electric	284	13.1%
	Petrol	4	0.2%
Site No.	1	547	25.2%
	2	171	7.9%
	3	666	30.7%
	4	783	36.1%

Analysis

An LCA approach was employed in this study to identify latent behavioral subgroups of two- and three-wheel vehicle users. LCA is a finite mixture model which attempts to group individuals into different homogeneous classes based on a latent class variable that explains the groupings through tested observed indicators. Chosen covariates are used as a means of characterizing the different identified classes and predicting class membership. Covariates included in the model are sociodemographic characteristics, vehicle characteristics, and environmental characteristics. The LCA calculations are created using LatentGOLD (version 6.1) software (Vermunt and Magidson, 2016).

The number of classes selected for LCA were chosen in consideration of the identification criteria, BIC, AIC, and CAIC (Collins & Lanza, 2009). Table 2 presents the metrics used for comparison of relative model fit and selection. It is generally considered for LCA that a model with a lower BIC value represents a better balance between model fit and parsimony. Therefore, based on the estimated identification criteria, a model with three classes was considered the most appropriate. Further, the entropy value of 0.92 for a three-class model indicates a good separation between the classes identified (Celeux & Soromenho, 1996).

Table 2. Model fit statistics and class identification criteria

Classes	Parameters	LL	AIC	BIC	CAIC	Entropy R^2
2	64	-8455.62	17039.25	17402.84	17466.84	0.91
3	**113**	**-8207.38**	**16640.77**	**17282.74**	**17395.74**	**0.92**
4	162	-8097.88	16519.76	17440.10	17602.10	0.82
5	211	-8018.91	16459.82	17658.53	17869.53	0.90

3. RESULTS AND DISCUSSION

The profile of the indicators and active covariates present in the final LCA are presented in Table 3. While it is generally considered that a class assignment which proportionally accounts for less than 5% of the sample could indicate instability within the class itself, the profile of the observed Class 3 represents a class exhibiting extreme behavior not captured within the other observed classes. Based on this, as well as the fact that such distracted behaviors which are present in the model only account for a small proportion of the overall sample, it is reasonable to consider that this observed Class 3 represents a behaviorally distinct class of two- and three-wheel vehicle users exhibiting little instability.

Table 3. LCA model profile for indicators and covariates

		Class 1	Class 2	Class 3
	Class size	50.0%	47.9%	2.1%
Indicator				
Steering control	**Two hand steering**	91.6%	91.6%	11.4%
	One hand steering	8.3%	7.7%	86.4%
	No hand steering	0.1%	0.7%	2.2%
Control obstruction	**No**	97.9%	97.9%	100.0%
	Yes	2.1%	2.1%	0.0%
Headphone use	**No**	67.3%	69.5%	67.4%
	Yes	32.0%	29.4%	32.6%
	Unknown	0.7%	1.1%	0.0%
Phone use	**No**	98.3%	99.7%	22.3%
	Mounted phone use	1.3%	0.3%	0.0%
	Phone in hand	0.5%	0.0%	77.7%
Helmet use	**No helmet use**	66.1%	63.9%	89.0%
	Helmet present, in use	33.1%	36.0%	11.0%
	Helmet present, not in use	0.7%	0.1%	0.0%
Normal Platooning	**No**	2.6%	100.0%	63.0%
	Yes	97.0%	0.0%	37.1%
	Unknown	0.4%	0.0%	0.0%
Social Platooning	**No**	98.1%	92.2%	100.0%
	Yes	1.9%	7.8%	0.0%
Traffic in bike lane	**Single vehicle in bike lane**	7.4%	55.3%	69.3%

	≤2 vehicles in bike lane	44.1%	39.3%	24.0%
	≤4 vehicles in bike lane	24.7%	5.4%	0.1%
	>4 vehicles in bike lane	23.8%	0.0%	6.6%
Covariate				
Rider age	**1-12 yrs**	0.6%	1.8%	8.7%
	13-18 yrs	4.0%	4.5%	0.0%
	19-24 yrs	17.3%	18.5%	28.3%
	25-44 yrs	58.2%	53.9%	63.0%
	45-64 yrs	16.5%	17.1%	0.0%
	>65 yrs	3.5%	4.2%	0.0%
Rider gender	**Male**	51.0%	53.6%	47.8%
	Female	49.0%	46.4%	52.2%
Handlebar configuration	**Flat bar**	32.8%	35.8%	37.0%
	Flat raise bar	29.4%	28.0%	28.3%
	Drop bar	6.5%	7.4%	0.0%
	Cruiser	31.1%	28.5%	34.8%
	Bull horn bar	0.2%	0.3%	0.0%
Brake configuration	**Right- and left-hand brake**	40.0%	44.0%	23.9%
	Right-hand brake only	10.1%	9.4%	15.2%
	Left-hand brake only	49.9%	46.7%	60.9%
Vehicle design	**City bike**	87.5%	84.1%	87.0%
	Road bike	5.6%	7.6%	0.0%
	Mountain bike	2.0%	2.0%	2.2%
	Cargo bicycle	4.5%	5.6%	10.9%
	Scooter	0.4%	0.3%	0.0%
	Motorcycle (Moped)	0.0%	0.4%	0.0%
Vehicle function	**Personal**	89.4%	93.4%	89.1%
	Extended rental	3.8%	1.6%	0.0%
	Shared	6.4%	5.0%	10.9%
	Taxi	0.4%	0.0%	0.0%
	Delivery	0.1%	0.0%	0.0%
Vehicle power	**None**	87.5%	86.5%	73.9%
	Electric	12.4%	13.3%	26.1%
	Petrol	0.2%	0.2%	0.0%
Time of day	**05:00-0515**	0.1%	1.2%	0.0%
	07:00-07:15	16.1%	15.5%	6.5%
	09:00-09:15	20.3%	14.3%	8.7%
	11:00-11:15	4.9%	11.4%	19.6%
	13:00-13:15	5.9%	14.6%	21.7%
	15:00-15:15	18.6%	15.4%	21.7%
	17:00-17:15	27.0%	17.0%	2.2%
	19:00-19:15	7.0%	10.7%	19.6%
Site No.	**1**	20.2%	31.1%	13.0%
	2	2.9%	12.5%	21.7%
	3	31.0%	30.9%	21.7%
	4	45.9%	25.6%	43.5%

Class 1 can be described as platooned riders in high volume/peak hour traffic and might be closely defined as commuters. This distinct behavioral group of two- and three-wheel vehicle users accounts for most riders sampled, making up 50% of the data used for analysis. This group is made up almost entirely (97%) of platooned riders operating in high volume traffic. Notably, the highest proportion of riders belonging to this class assignment fall within morning peak and afternoon peak hours and at both of Sites 3 and 4, the high interaction sites. Therefore, this group exhibits high exposure regarding level of traffic and potential interaction with pedestrians.

Class 2 is another relatively large grouping representing independent and social riders in low traffic, accounting for approximately 48% of the sampled observations. The individuals which belong to class 2 were found to operate mostly in lower to no traffic and are more often platooning socially (8%) alongside another rider than their class counterparts. Moreover, this class has a higher proportion of riders at Site 1, a non-interaction site. Indeed, both class 1 and class 2 are quite intuitive and largely reflect expected traffic dynamics (e.g. high traffic leading to platooning, low traffic enabling more independent riding).

The final class, class 3, consists of distracted/risk accepting riders and is proportionally much smaller than its counter parts, only making up approximately 2% of the sampled data. Indeed, this group represents the individuals using their phones while riding (77%) and riding without full control (one hand steering) of their vehicle (86%). Further, this group is also riding in no traffic (69%) and without helmets (89%). Individuals which fall into class 3 assignment are also mostly riding in afternoon and evening hours with more of them riding electric powered vehicles (26%) than in the other groups. Therefore, this group represents a risk accepting behavioral group, as they are engaging in multiple distracted and safety critical behaviors.

Overall, the identified classes suggest that rider behavior is primarily structured by traffic context and moment-to-moment interactions rather than by stable vehicle-related characteristics. The two largest groups largely reflect expected traffic dynamics, where higher density conditions are associated with more constrained, platooned riding, and lower density conditions allow for more independent behavior. In this sense, these classes provide a useful validation of the observational approach and the LCA in capturing meaningful patterns in the data. At the same time, the identification of a smaller group characterized by distraction and reduced control highlights how specific combinations of behaviors and contexts may concentrate higher-risk patterns, which are not visible when treating riders as a homogeneous population.

4. CONCLUSIONS

The aim of this study was to identify latent behavioral groups among two- and three-wheel vehicle users in Copenhagen using naturalistic observational data through an LCA approach. Results demonstrate that the sampled riders are not homogenous but rather exhibit distinct behavioral classes that are characterized by different traffic interaction patterns, traffic exposure, distracted, safety critical variables, and environmental characteristics.

This research demonstrates the value in naturalistic observational studies for uncovering latent behavioral patterns for providing new insights and understanding into a two- and three-wheel vehicle safety. Further, to address the behavioral groups identified in this study targeted safety measures should be developed. For

instance, to limit exposure and platooning of riders in class 1, traffic calming measures should be taken e.g. optimized signal timings for riders during peak hours and lane markings to improve lateral positioning of riders in the bike lane for safer overtaking. For class 2, awareness programs and signage promoting situational awareness might limit risk from overtaking and improve awareness when riders are social platooning. Finally, to address distracted riders (class 3), there should be higher policy enforcement as it is forbidden to use a mobile phone while riding in the bike lane in Denmark. Additionally, awareness campaigns on vehicle control and phones use could target younger individuals and would make this group of riders aware of the risks of distracted riding.

REFERENCES

Amado, S., Arıkan, E., Kaça, G., Koyuncu, M., & Turkan, B. N. (2014). How accurately do drivers evaluate their own driving behavior? An on-road observational study. Accident Analysis & Prevention, 63, 65-73.

Aydar, R., Çelik, O. N., & Afşar, A. (2025). The impact of traffic infrastructure, environment, and user behavior on urban bicycle use: a literature review. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, *31*(5), 849-864.

Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. Journal of classification, 13(2), 195-212.

Collins, L. M., & Lanza, S. T. (2013). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. John Wiley & Sons.

Damant-Sirois, G., Grimsrud, M., & El-Geneidy, A. M. (2014). What's your type: A multidimensional cyclist typology. *Transportation*, *41*(6), 1153-1169.

de Freitas, L. M., & Axhausen, K. W. (2025). What influences cycling infrastructure preferences? A stated-preference survey. *Transportation*, 1-20.

Dill, J., & McNeil, N. (2013). Four types of cyclists? Examination of typology for better understanding of bicycling behavior and potential. *Transportation Research Record*, *2387*(1), 129-138.

Eby, D. W. (2011). Naturalistic observational field techniques for traffic psychology research. In *Handbook of traffic psychology* (pp. 61-72). Academic Press.

Fisa, R., Musukuma, M., Sampa, M., Musonda, P., & Young, T. (2022). Effects of interventions for preventing road traffic crashes: an overview of systematic reviews. *BMC public health*, *22*(1), 513.

Friedman, W. J. (1993). Memory for the time of past events. *Psychological bulletin*, *113*(1), 44.

Grimm, P. (2010). Social desirability bias. *Wiley international encyclopedia of marketing*.

Griswold, J. B., Yu, M., Filingeri, V., Grembek, O., & Walker, J. L. (2018). A behavioral modeling approach to bicycle level of service. *Transportation research part A: policy and practice*, *116*, 166-177.

Nguyen-Phuoc, D. Q., Oviedo-Trespalacios, O., Nguyen, T., & Su, D. N. (2020). The effects of unhealthy lifestyle behaviours on risky riding behaviours–A study on app-based motorcycle taxi riders in Vietnam. *Journal of Transport & Health*, *16*, 100666.

O'Hern, S., Stephens, A. N., Siebert, F. W., & Useche, S. A. (2026). Is our growing affinity for technology a challenge for preventing distracted cycling? An Australian study. *Traffic injury prevention*, *27*(2), 180-187.

Ramos-Ruiz, J. E., Muñoz-Fernández, G. A., García García, L., & Arteaga Sánchez, R. (2025). Pedalling towards satisfaction: the power of motivation and emotions in shaping tourism perception in cycle tourism events. *Journal of Sport & Tourism*, *29*(2), 107-131.

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

Siebert, F. W., Ringhand, M., Englert, F., Hoffknecht, M., Edwards, T., & Rötting, M. (2021). Braking bad–Ergonomic design and implications for the safe use of shared E-scooters. Safety science, 140, 105294.

Useche, S. A., Alonso, F., Montoro, L., & Esteban, C. (2018). Distraction of cyclists: how does it influence their risky behaviors and traffic crashes?. *PeerJ*, *6*, e5616.

Vermunt, J. K., & Magidson, J. (2005). Latent GOLD 4.0 user's guide.