Cabral, Laura, and Kim, A.

An empirical reappraisal of the four types of cyclists.

AUTHOR POST PRINT VERSION

Cabral, L., & Kim, A. (2020). An empirical reappraisal of the four types of cyclists.
Transportation Research Part A: Policy and Practice, 137, 206–221.

https://doi.org/10.1016/j.tra.2020.05.006

An Empirical Reappraisal of the Four Types of Cyclists

Laura Cabral, M.Sc., Corresponding Author
Department of Civil and Environmental Engineering
6-362 Donadeo Innovation Centre for Engineering, University of Alberta, Edmonton, AB, Canada, T6G 1H9
Phone: 819-993-1901
Email: lcabral@ualberta.ca
ORCID: 0000-0002-9553-7606

Amy M. Kim, Ph.D.
Associate Professor
Department of Civil and Environmental Engineering
6-269 Donadeo Innovation Centre for Engineering, University of Alberta, Edmonton, AB, Canada, T6G 1H9
Phone: 780-492-9203
Email: amy.kim@ualberta.ca
ORCID: 0000-0002-5250-2796

Declarations of interest: none.

ABSTRACT

The Four Types of Cyclists is a widely adopted typology developed by Portland's Bicycle Coordinator, Roger Geller. No Way No How, Interested but Concerned, Enthused and Confident, and Strong and Fearless have become ubiquitous in academic literature and practice. However, the classification was subjectively developed and contains several known contradictions. This research aims to develop a data-driven typology using near-identical explanatory variables to those of the Four Types of Cyclists. The objective is to develop a typology with a similar functional purpose, but derived using statistical methods. An online survey was distributed to a panel of Edmonton, Canada, residents to this effect, and the use of video clips rather than descriptions is tested as a means to assess comfort on different types of cycling infrastructure. Cluster Correspondence Analysis is used to carry out the segmentations, including variables of comfort, cycling intent, and cycling in the previous summer. The survey sample tends to segment into three categories, as opposed to the four suggested by Geller: Uncomfortable or Uninterested, Cautious Majority, and Very Comfortable Cyclists. The Four Types of Cyclists typology is also shown to generate heterogeneous comfort patterns within each cyclist type, a limitation our empirically-derived segmentations overcome.

Keywords: Cyclist segmentation, cyclist type, typology, comfort, intent, bicycle facilities.

1 INTRODUCTION

The Four Types of Cyclists typology is a popular cyclist classification developed by Portland's Bicycle Coordinator, Roger Geller, in 2006. Comprised of No Way No How, Interested but Concerned, Enthused and Confident, and Strong and Fearless, it has become a widely adopted, and adapted, method for classifying cyclists and potential cyclists, both in research and in practice (Dill and McNeil, 2013; Félix et al., 2017). The objective in segmenting the cycling (or non-cycling) population is often to determine appropriate policies to accommodate or encourage cycling for different types of (potential) cyclists (Félix et al., 2017). This was in part Roger Geller's objective: to understand Portland's market for cycling based on existing surveys and expert knowledge (Geller, 2006).

Dill and McNeil (2013) later formalized a method to classify a population into the Four Types of Cyclists and reaffirmed the typology using a wide sample of American cities (Dill and McNeil, 2016). The variables used to derive cyclist types include stated comfort on different types of infrastructure, the intent to cycle more often, and use of a bicycle in the last 30 days. The methodology developed by Dill and McNeil (2013) uses a rule-based approach to determine the cyclist type. The ease of application and the intuitiveness of the typology have led to its adoption for other important applications.

Notably it was the basis for a popular low-stress cycling network connectivity assessment method: the Level of Traffic Stress (LTS) (Mekuria et al., 2012). The framework classifies streets and other cyclable environments based on infrastructure characteristics and their expected effect on cyclist stress. The four LTS levels roughly map onto the Four Types of Cyclists, with LTS 2 identifying cycling environments that are suitable for the Interested but Concerned, LTS 3, the Enthused and Confident, and LTS 4, the Strong and Fearless. The LTS classification has also been widely adopted in various locales to assess connectivity, prioritize future projects, and analyze collision patterns in safety analyses, (Cabral et al., 2019; Chen et al., 2017; Kent and Karner, 2018; Moran et al., 2018; Semler et al., 2018) and, by extension, has made the Four Types of Cyclists ubiquitous in many areas of cycling research and practice. Note that, despite their wide adoption, no research has empirically demonstrated a correspondence between the Four Types and their respective intended LTS level.

Despite its popularity, the typology is not without flaws. First, the segmentation was subjectively developed based on expert knowledge (Geller, 2006), therefore the categories are imposed on the population, rather than being empirically developed from it. Second, some unanticipated contradictions emerge. For example, the assumed relationship of higher cycling frequency with higher comfort does not hold throughout the literature; in particular many Strong and Fearless respondents are not, in fact, cyclists, as pointed out by Damant-Sirois and El-Geneidy (2015). This may be linked to the subjective development of the typology or could also be an artefact of the rule-based method developed by Dill and McNeil (2013).

This research was prompted by the above issues, as well as previous research results where we found cyclists to naturally form three categories rather than four when assessed using comfort variables alone (Cabral et al., 2018). Although and because these categorizations have been so widely adopted, more empirical testing and validation is warranted. We propose a reappraisal of the Four Types of Cyclists based on an empirical segmentation method developed with data collected in Edmonton, Canada. Our approach is to use variables as similar as possible to the original typology, such that any new segmentation that results can serve a similar purpose as the Four Types of Cyclists. In particular, we aim to make our typology amenable to an LTS-like connectivity assessment framework with the objective of forming groups of cyclists of similar comfort level in a more consistent way.

After presenting the relevant literature, we detail our method to empirically develop a cyclist typology from similar inputs. The analysis and discussion compare the classic Four Types of Cyclists typology with two empirically obtained typologies, one developed using text descriptions of cycling facilities, and one using videos. A secondary objective of this work is to compare the suitability of using text descriptions versus videos in the definition of cyclist types. Finally, the conclusions touch on implications for the field.

2 LITERATURE REVIEW

We first provide an overview of several existing cyclist segmentations, their purpose, and the methods to derive them. We then review the relevance of the variables used to define the Four Types of Cyclists.

2.1 Cyclist Typologies

It is well recognized that cyclists are a heterogeneous group; policies impact cyclists differently depending on their particular characteristics (Damant-Sirois et al., 2014; Kroesen and Handy, 2014), and their route and infrastructure preferences vary accordingly (Larsen and El-Geneidy, 2011; Stinson and Bhat, 2005; Veillette et al., 2019). This explains the relative abundance of segmentations found in the literature attempting to categorize cyclists. One of the older and well-known classifications is the ABC typology adopted by the United States Federal Highway Administration (Advanced bicyclists, Basic bicyclists, Children) (Wilkinson et al., 1994).

Félix et al. (2017) offer an excellent review of different cyclist segmentations from peer-reviewed and grey literature published between 1994 and 2014. Common reasons to create cyclist typologies are planning infrastructure and cycling policies, understanding sociodemographic profiles of different types of cyclists, or for the sake of segmenting *per se*. Variables used often include frequency and purpose of cycling and variations on the theme of comfort, cycling confidence, risk perception, and experience. A few typologies also include socio-demographic variables.

Two broad methodological categories are identified by Félix et al. (2017): bottom-up and top-down. The first category refers to empirical segmentations, usually found in the academic literature and involving the use of factor analysis, clustering methods, or a combination of both to derive cyclist types from data. The second is more commonly found in the grey literature (although not exclusively) and relies on expert knowledge to define cyclist types. Like the Four Types of Cyclists, these sometimes use rule-based methods to classify respondents.

One method borrowed from market segmentation research has been adopted in several empirical cyclist typology development papers: a principal component analysis (PCA) followed by clustering (usually k-means). The PCA is first carried out using a series of variables of interest for the segmentation. This allows a reduction from a large set of variables to a limited number of underlying factors. Clustering is then applied to the reduced set of factors to yield a given number of cyclist types with similar characteristics. Using this method, different typologies can be derived based on the characteristics of interest for the researchers. For example, Damant-Sirois et al. (2014) defined four types of cyclists and then used the classification to measure the importance of different factors in determining the cycling frequency for each type (Damant-Sirois and El-Geneidy, 2015). The same segmentation methodology was used by Gatersleben and Haddad (2010) to define cyclist stereotypes. Their aim was not to segment cyclists *per se*, but rather to understand how cyclists are perceived by both cyclists and non-cyclists.

Two segmentations published more recently are noteworthy since their aim is in line with ours: defining cyclist types to uncover infrastructure preferences. Veillette et al. (2019) used the tandem PCA and k-means clustering method to define six types of cyclists. Using 29 variables grouped into 9 factors by PCA, their final typology after k-means clustering of factors includes urban, benefit-seeking, happy, picky-efficiency, childhood-influenced and indifferent cyclists. Their objective was to understand how the use of three types of facilities (recreational paths, bidirectional protected bike lanes, and painted lanes) was influenced by membership in each of the six categories. The purpose was to provide nuanced infrastructure preference information for each cyclist type to inform future infrastructure planning.

Griswold et al. (2018) sought to develop a new quantitative bicycle LOS measure, using latent class choice models to segment cyclists and uncover their infrastructure preferences in one integrated framework. The latent class model defined three types of cyclists based on demographics, experience, and cycling preferences: neighborhood, urban, and fitness cyclists. The class-specific models identify the specific infrastructure preferences for each cyclist type. The results were obtained from convenience sampling of a limited number of respondents (221 online and 14 in person); this means the three types defined should be considered with caution. Nonetheless, their work highlights that the three types are not on an ordinal scale of comfort, contrarily to the Four Types of Cyclists, and the authors argue that cyclist typologies developed specifically to help with infrastructure choice do not necessarily have to follow an ordinal pattern.

2.2 Relevant Segmentation Variables

While some typologies seek to integrate as many of the known determinants of cycling as possible to create a holistic segmentation (e.g. (Damant-Sirois et al., 2014)), the Four Types of Cyclists rely on three core variables: comfort perception of different types of cycling environments, intent to cycle more often than one currently does, and frequency of cycling in the recent past. Because our purpose is not to modify the basis of the typology, we briefly discuss the relevance of these three variables.

2.2.1 Comfort

A lack of perceived and actual comfort and safety is associated with lower levels of cycling (Dill and Voros, 2007; Winters et al., 2011) and is a known, major deterrent to cycling (Heinen et al., 2010; Lois et al., 2015). Comfort can be affected by many variables, including facility type and width, physical separation from traffic, presence and occupancy of parking, land use type, etc. (Blanc and Figliozzi, 2016; Heinen et al., 2010; Li et al., 2012). Amongst other demographics, gender influences the perception of safety, with women more likely to perceive their cycling environment as more dangerous or less comfortable than men (Sener et al., 2009).

2.2.2 Intent

Intent quantifies the pool of potential future cyclists as it is a predictor of actual behaviour (Ajzen, 1991), even though an intention-behaviour gap is often observed (Sheeran, 2002). Intent to cycle can be influenced by social and self-identity (Heinen, 2016; Lois et al., 2015), subjective norms (i.e. how cycling is perceived by the respondent's peers) and descriptive norms (i.e. whether others in the respondent's circle cycle) (Eriksson and Forward, 2011). Self-efficacy (the knowledge and skills required to ride and maintain a bicycle) and current cycling frequency are also found to influence intent (Lois et al., 2015). Finally, Cabral et al. (2018) found no statistically significant difference in cycling intent between respondents of different genders.

2.2.3 Frequency

To determine membership within the Four Types of Cyclists, Dill and McNeil (2013) include a variable indicating whether a respondent cycled in the last 30 days, which helps to differentiate between the No Way No How and Interested but Concerned membership. The purpose of including this variable is to distinguish cyclists from non-cyclists and, a distinction that is often included in cyclist segmentations (Félix et al., 2017). Damant-Sirois and El-Geneidy (2015) found that perceived safety throughout the network, the low cost of cycling, and its perceived convenience are associated with higher cycling frequency. It is also associated with cycling for multiple purposes in addition to commuting (e.g. recreation, running errands, etc.) (Stinson and Bhat, 2004), which seems to be a gateway to bicycle commuting (Sener et al., 2009). As was the case for comfort, gender influences cycling frequency with men generally cycling more than women (Cabral et al., 2018; Heinen et al., 2010; Sener et al., 2009).

Overall, the literature shows the three variables included in the Four Types of Cyclists are pillar characteristics of cycling and are highly interrelated.

3 METHODOLOGY

To assess the suitability of the Four Types of Cyclists, we first needed to reproduce the method described in Dill and McNeil (2013) as a comparison point to the empirical segmentation. To this effect, we developed and distributed a survey in fall 2018, as described in Data (3.1); the analytical method is described in 3.2.

3.1 Data

Our data was collected using the Bicycle Ridership and Traffic Stress Tolerance survey, which we developed to assess multiple aspects of cyclist comfort. The survey was distributed in collaboration with the City of Edmonton, using their in-house survey platform and the Insight Community, a panel of Edmontonians who sign up to answer surveys from the City regularly. In addition, to reach a larger group and obtain a broader range of perspectives, an open link to the survey was distributed through university mailing lists, local active transportation advocacy groups, community league emails, and media outlets. Both cyclists and non-cyclists were encouraged to participate. After data cleaning, 3208 valid responses remained, with 2193 responses from Insight Community members (24% response rate) and 1015 from other Edmonton residents. Since the survey sample is not random, it is worth comparing key demographics from survey respondents to data for the City of Edmonton obtained from the 2016 Canadian census (Statistics Canada, 2017). Table 1 shows that the survey sample and population data are statistically significantly different regarding age, income, and education. In terms of gender, only the proportion of male respondents is equivalent to the proportion of men in the census data. Notable differences include an underrepresentation of 15 to 24 year-olds, and a large overrepresentation of university-educated respondents in the sample. Since 16.4% of respondents did not answer the income question, the comparison to the census data is less straightforward, but there appears to be a sizable underrepresentation of respondents in households where the total income is below $50,000. These differences were expected based on previous use of the Insight Community data (Cabral et al., 2018) and given the format and channels of distribution of the survey.

Table 1 Comparison of Key Demographic Characteristics Between the Survey Sample and the City of Edmonton Population

Variable	Proportion of Survey Respondents (n = 3208)	Proportion from Edmonton Census Data (n = 932,546)
Gender		
Female	45.4	50.0
Male	51.4	50.0
*Other**	3.2	-
Age (Years) †		
15 – 24	4.1	15.6
25 – 44	47.0	39.8
45 – 64	35.9	29.9
65 and over	12.2	14.6
Prefer not to answer	0.8	-
Income ($ CAD)		
< 50,000	9.9	25.6
50,000 to 99,999	26.3	31.7
100,000 or more	47.4	42.6
Prefer not to answer	16.4	-
Education		
High school or less	9.6	42.8
Technical school	22.4	29.8
University degree	65.7	27.3
Prefer not to answer	3.3	-

Note: All variables are statistically significantly different (($p < 0.05$), except the proportion of males.
* Includes “Prefer not to answer” and “Neither describes me” in our survey.
† Census data proportions adjusted to exclude 14 years old and younger categories.

The survey elements used to carry out the empirical segmentations are listed in Table 2 and Table 3[1]. As noted above, our first aim was to reproduce the Four Types of Cyclists segmentation based on the established methodology (Dill and McNeil, 2013, 2016). The list of infrastructure descriptions (*Comfort variables*, Table 2) was adapted to our local Canadian context and edited for conciseness. For example, speeds were converted to kilometers per hour (kph), using commonly found speed limits rather than an exact conversion. Some redundant or underutilized descriptions were also eliminated and cycling frequency was measured by season. These changes to the original survey limit the comparability of the resulting cyclist classification, but are necessary to collect meaningful answers from Canadian respondents. The elimination of redundant descriptions was also deemed reasonable to avoid survey fatigue. Indeed, in addition to rating

[1] All questions other than demographics were mandatory and respondents could not progress through the survey without answering the questions. Two responses were received with incomplete information for mandatory fields, likely due to a survey malfunction. These responses were eliminated.

descriptions, respondents were asked about their general cycling behaviour (e.g. frequency, purpose, also shown in Table 2), their demographic characteristics, and had to rate video clips, as described below.

Table 2 Segmentation Variables (Excluding Videos, see Table 3)

Variable	Statement	Possible Responses
Comfort	How comfortable would you feel riding in these different environments?	
T_Path	A path or trail separated from the street.	
T_Quiet_Residential	A quiet residential street with low traffic speeds.	
T_Sharrow	A quiet residential street with bike route signs and shared-use lane or sharrow markings.	
T_Local_Commercial	A neighborhood commercial shopping street with one lane in each direction, traffic speeds of 40 to 50 km/hour, on-street car parking, and no reserved (painted) bike lane.	Very Uncomfortable Somewhat Uncomfortable Somewhat Comfortable Very Comfortable
T_Local_Commercial_BL[1]	A neighborhood commercial shopping street with one lane in each direction, traffic speeds of 40 to 50 km/hour, on-street car parking, and a reserved (painted) bike lane.	
T_Major	A major street with two lanes in each direction, on-street parking, traffic speeds of 50 to 60 km/hour and no reserved (painted) bike lane.	
T_Major_BL[1]	A major street with two lanes in each direction, on-street parking, traffic speeds of 50 to 60 km/hour and a reserved (painted) bike lane.	
T_Major_PBL[2]	A major street with two lanes in each direction, on-street parking, traffic speeds of 50 to 60 km/hour and a bike lane physically protected from traffic by bollards or planters.	
Intent	Please indicate your level of agreement or disagreement with the following statement: I would like to travel by bike more than I do now.	Strongly Disagree Somewhat Disagree Neither Agree nor Disagree Somewhat Agree Strongly Agree
Frequency Freq_Fall16 Freq_Winter17 Freq_Spring17 Freq_Summer17 Freq_Fall17 Freq_Winter18 Freq_Spring18 Freq_Summer18	In the last year (Sept. 2017 to August 2018) how often did you ride a bike during each season? Please indicate your cycling frequency only if you lived in Edmonton during each period. Think back to the previous year (Sept. 2016 to August 2017). How often did you ride a bike during each season? Please indicate your cycling frequency only if you lived in Edmonton during each period.	Daily 4 or more times per week 2 to 3 times per week Once per week Once per month Never
Biked_Summer18	Binary variable created from *Freq_Summer18*. Indicates if respondent biked at least once a month or not in Summer 2018.	Yes No
Purpose	In the last year (Sept. 2017 to August 2018), for which reasons did you ride your bike? Select all that apply.	Recreation Fitness Utility Commute

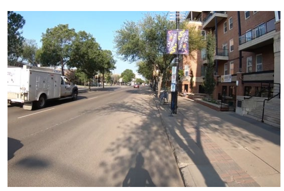
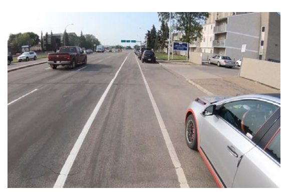
[1] BL stands for "Bicycle Lane"
[2] PBL stands for "Protected Bicycle Lane"

The survey included video clips (8-12s each) of various local cycling environments that aimed to offer a more immersive tool to glean cyclist comfort (Table 3). Research has previously shown respondents can discern most roadway environment conditions as accurately with videos as if they were in the field (Harkey et al., 1998) making them an effective assessment tool (Harkey et al., 1998; Jensen, 2007; Landis et al., 1997; Lehtonen et al., 2016; Parkin et al., 2007), while also minimizing risk and allowing extensive data collection, including from participants who do not cycle.

The video footage was collected by the research team while riding their bikes at different locations throughout Edmonton. This was done using a GoPro Hero 6 camera, which allowed smooth recording with no post-processing. Of the 26 locations initially filmed, some were removed due to poor video quality and others, to avoid survey fatigue. The remaining 16 were retained and presented to respondents. Eight videos were shot in locations as similar as possible to the eight infrastructure descriptions used in the Four Types of Cyclists segmentation (*Comfort* variables, Table 2), within the constraints of available infrastructure in Edmonton, and traffic conditions at the time of recording. These eight equivalent videos are used for the analysis presented in this paper and are shown in Table 3. The remaining eight videos were added to the survey to capture a wider variety of cycling environments available in Edmonton, and are used for further exploratory segmentation beyond the scope of this work (Cabral, 2019). A more detailed description of each video is available in (Cabral, 2019).

Although videos have become ubiquitous in survey research, the specific study of differences in survey responses between audiovisual stimuli and written descriptions has only been explored in a handful of studies (Shaw et al., 1992; Sleed et al., 2002); none pertain to the transportation field at large. Generally, videos have complete contextual information, while written descriptions focus on a limited set of variables. With videos, respondents have the same contextual cues, and unspecified aspects are not left to the imagination (Sleed et al., 2002) while also limiting influence from lack of prior knowledge about the subject (Shaw et al., 1992). On the other hand, the rich context does not allow to control which variables the respondents focus on, contrary to the written descriptions. While more research is needed, videos are found to portray the given situation in a more realistic way than the written equivalent (Smith and Sokolowski, 2008).

Table 3 Equivalent Video Clips to Text Descriptions* Used for Segmentation

Variable	Equivalent Text Variable	Representative Frame	Possible Responses
V_Path	T_Path		
V_Quiet_Residential	T_Quiet_Residential		
V_Sharrow	T_Sharrow		
V_Local_Commercial	T_Local_Commercial		Very Uncomfortable Somewhat Uncomfortable Somewhat Comfortable Very Comfortable
V_Local_Commercial_BL[1]	T_Local_Commercial_BL[1]		
V_Major	T_Major		
V_Major_BL[1]	T_Major_BL[1]		
V_Major_PBL[2]	T_Major_PBL[2]		

* Equivalent to the eight facility descriptions presented in Table 2 (Comfort variables).
[1] BL stands for "Bicycle Lane"
[2] PBL stands for "Protected Bicycle Lane"

3.2 Analysis

Given its popularity as an empirical cyclist segmentation method (Damant-Sirois et al., 2014; Gatersleben and Haddad, 2010; Veillette et al., 2019), we initially envisioned using tandem PCA and k-means clustering for this work. However, the method requires at least a five-point Likert-type scale such that responses can be assimilated to continuous data (Markos et al., 2018). Our comfort ratings emulate Dill and McNeil's (2013) who required respondents to rate infrastructure descriptions on a four-point scale. These responses can hardly be assimilated to continuous data and should be treated as categorical. An equivalent PCA method for categorical variables is Multiple Component Analysis (MCA), which can be followed by k-means clustering to obtain similar segmentation results (Markos et al., 2018). However, it has been shown that this tandem approach, despite its popularity, does not yield the best possible results as the PCA/MCA and k-means clustering optimize different functions and the original dimension reduction can mask crucial variables to identify niche segments (Dolnicar and Grün, 2008; van de Velden et al., 2017). van de Velden et al. (2017) proposed Cluster Correspondence Analysis (CCA), a method that simultaneously reduces and clusters categorical data to create the segmentation. This method identifies cluster membership and category weights such that between cluster and between category variances are simultaneously maximized. Cluster membership and variable categories are cross-tabulated to this effect, allowing a single objective function to be maximized, thus eliminating the main drawback of the tandem approach. The mathematical details of the method are presented in van de Velden et al. (2017).

CCA has been used to segment different populations into categories with similar characteristics. For example, in their demonstration of the R package *clustrd* (Markos et al., 2019), which implements CCA, Markos et al. (2018) identified groups of similar Indonesian women with respect to socio-economic factors and choice of contraceptive method. CCA results have also been used to make policy recommendations. This was the case in a study of student profiles (Papageorgiou et al., 2016), which aided in the identification of potential gaps in teaching instruments used to discuss the atom, and in the formulation of recommendations to address those gaps.

We use CCA, implemented in the R package *clustrd* (Markos et al., 2019), to derive the empirical segmentations using variables as close as possible to those used to determine membership in each of the Four Types of Cyclists. We explore two different segmentations:

1. **'Text Description' Segmentation**
   This segmentation is meant to be as close as possible to the Four Types of Cyclists as it includes the same input variables: the eight Facility variables, Intent, and Biked_Summer18.

2. **'Video Equivalent' Segmentation**
   This segmentation is the same as the 'Text Description' segmentation, but the eight Facility variables are replaced by the eight equivalent video clips listed in Table 3. We expect this segmentation to yield a more representative classification given video clips offer a more immersive experience, and provide the same contextual cues to all respondents, regardless of cycling experience.

To determine the most appropriate number of clusters and dimensions for each segmentation, we use several diagnostic tools. The *tuneclus* function in the *clustrd* package assesses average silhouette width (ASW, Rousseeuw (1987)), with higher values indicating better separation and more compact clusters. We also employed bootstrapping methods to assess cluster stability. The *global_bootclus* function iterates through numbers of clusters and returns adjusted Rand index (ARI) values, with ARI > 0.8 considered an indication of good stability (Steinley, 2004). Finally, the *local_bootclus* function also uses a bootstrapping method to assess the stability of each cluster within a given solution with *m* clusters; a Jaccard index above 0.75 indicates a stable cluster (Hennig, 2007). We determine an appropriate number of clusters for each segmentation using these diagnostic tools and considering the interpretability of the resulting clusters. For each set of variables, we evaluated solutions with three to eight clusters. The number of dimensions can range from 1 to *m*–1, where *m* is the number of clusters.

To understand the distinctive attributes of each cluster, we examine variable-categories (e.g. Intent = *strongly disagree*) with high standardized residuals. High values indicate a significantly higher observed frequency for a given variable-category in the cluster, compared with the observed frequency for all survey respondents. For large samples such as ours, the standardized residual values can be interpreted in a similar way to z-scores. Values of two and above are considered significant at the 95% level, and three and above, at the 99% level.

4 FINDINGS AND COMPARISONS TO THE EXISTING CLASSIFICIATION

We first present the distribution of respondents that fall into the four established cyclist types and the groups resulting from the CCA analysis for the two empirical segmentations. Second, we compare and contrast comfort ratings between the three typologies, as well as demographic information. Finally, we examine the reclassification of respondents between the typologies to understand differences between the different approaches.

4.1 Definition of Groups within the Four Types of Cyclists and the Empirical Segmentations

We first classify survey respondents into the four types using Dill and McNeil's methodology (2013). Our sample is composed of 13.4% No Way No How, 70.3% Interested but Concerned, 11.1% Enthused and Confident, and 5.2% Strong and Fearless. Compared to results reported for 50 U.S. metropolitan areas (Dill and McNeil, 2016), our sample has a lower proportion of No Way No How and a higher proportion of Interested but Concerned. These differences can likely be attributed to our survey design, which did not use random sampling and thus, increased the likelihood of self-selection bias. In particular, some respondents who accessed the survey through the open link may have chosen to participate because of an underlying interest in the subject of cycling and some of the 24% of Insight Community members who accepted to participate may have done so for the same reason. Other factors may also contribute to differences in cyclist type distribution. Edmonton has an extensive network of paved recreational trails along the River Valley and connected ravines. The lower proportion of No Way No How cyclists may in part be explained by the high use of these trails, resulting in fewer respondents considering trails as *very uncomfortable* cycling environments. Finally, cycling culture is becoming more mainstream and interest may have increased in the years since Dill and McNeil (2016) collected their data. This may be particularly true for Edmonton, where significant active transportation investments (including bicycle infrastructure improvements and education campaigns) have taken place since 2017.

Table 4 shows test results on the clusters and sets of variable combinations considered for the two empirical segmentations. Although three to eight clusters were initially evaluated, the solutions with five to eight clusters were systematically of lower quality. In the interest of space, Table 4 only reports results for solutions with three or four clusters. In both cases, the ideal number of dimensions to maximize the ASW value was *m* - 1.

Table 4 Cluster Selection Statistics

Seg. Name	# of Clusters (*m*)	ASW	Adjusted Rand Index				Jaccard Index (median)			
			Min	*Max*	*Mean*	*Median*	*Clus. 1*	*Clus. 2*	*Clus. 3*	*Clus. 4*
Text Desc.	3	0.199	0.282	0.963	0.888	0.919	0.898	0.889	0.762	-
	4	0.161	0.424	0.953	0.878	0.909	0.846	0.789	0.826	0.520
Video Eq.	3	0.215	0.604	0.973	0.873	0.900	0.885	0.830	0.817	-
	4	0.181	0.659	0.960	0.867	0.884	0.881	0.800	0.911	0.609

From Table 4, we observe that the best solution contains three clusters for both the 'Text Description' and 'Video Equivalent' segmentations. Indeed, the highest ASW values are achieved with the three cluster solutions, which also have mean and median ARI values above 0.8, indicating good stability. Each of the three clusters in these solutions are also suitably stable. In contrast, the solutions with four clusters result in a fourth cluster below the 0.75 threshold in both cases (Jaccard = 0.520 and 0.614, respectively).

Table 5 shows the significant variable-categories for each cluster and both segmentations. These variable-categories define the salient characteristics of each cluster. On examination of these characteristics and those discussed in section 4.2, a proposed descriptive name is also offered for each cluster. Note that Table 5 indicates the number and proportion of respondents in each cluster; these proportions should not be necessarily considered representative of Edmonton's population given the biases of the sample previously discussed.

Table 5 Significant Variable-Categories – 'Text Description' and 'Video Equivalent' Segmentations.

Pale grey: significant at the 95% level (standardized residuals > 2). Dark grey: significant at the 99% level (standardized residuals > 3).

		'Text Description' Segmentation			'Video Equivalent' Segmentation		
Response Category	**Variable**[1]	Cluster #1: *Cautious Majority* (61.7%, n = 1979)	Cluster #2: *Comfortable Cyclists* (25.0%, n = 803)	Cluster #3: *Uncomfortable or Uninterested* (13.3%, n = 426)	Cluster #1: *Cautious Majority* (64.7%, n = 2075)	Cluster #2: *Very Comfortable Cyclists* (16.0%, n = 515)	Cluster #3: *Uncomfortable or Uninterested* (19.3%, n = 618)
Very Uncomfortable	_Path			Dark grey			Dark grey
	_Quiet_Residential			Dark grey			Dark grey
	_Sharrow			Dark grey			Dark grey
	_Local_Commercial			Dark grey			Dark grey
	_Local_Commercial_BL			Dark grey			Dark grey
	_Major			Pale grey			Dark grey
	_Major_BL			Dark grey			Dark grey
	_Major_PBL			Dark grey			Dark grey
Somewhat Uncomfortable	_Path			Pale grey			Pale grey
	_Quiet_Residential			Dark grey			Dark grey
	_Sharrow			Dark grey			Dark grey
	_Local_Commercial	Pale grey			Dark grey		
	_Local_Commercial_BL	Dark grey					
	_Major				Pale grey		
	_Major_BL	Dark grey			Pale grey		
	_Major_PBL						Dark grey
Somewhat Comfortable	_Path			Pale grey			Pale grey
	_Quiet_Residential	Pale grey					
	_Sharrow	Pale grey			Pale grey		
	_Local_Commercial		Dark grey		Pale grey		
	_Local_Commercial_BL	Dark grey			Dark grey		
	_Major		Dark grey				
	_Major_BL	Pale grey			Pale grey		
	_Major_PBL	Pale grey					
Very Comfortable	_Quiet_Residential		Dark grey			Dark grey	
	_Sharrow		Dark grey			Dark grey	
	_Local_Commercial		Dark grey			Dark grey	
	_Local_Commercial_BL		Dark grey			Dark grey	
	_Major		Dark grey			Dark grey	
	_Major_BL		Dark grey			Dark grey	
	_Major_PBL		Dark grey				
Strongly Disagree	Intent			Dark grey			Pale grey
Never	Biked_Summer18			Dark grey			Pale grey

Note: For increased legibility, variable-category combinations that are not significant in either segmentation are not included in the table.

[1] For table conciseness, text description variables and their equivalent video are presented on the same line.

For the 'Text Description' segmentation, the first cluster can be identified as the Cautious Majority. Cyclists or potential cyclists in this group are likely to be *somewhat comfortable* or *somewhat uncomfortable* on non-residential streets with and without painted bike lanes. The description of a protected bike lane (*T_Major_PBL*) is more likely to be rated as *somewhat comfortable* by this group than by the sample overall. The second largest cluster can be identified as Comfortable Cyclists, as they are more likely than the overall survey sample to find all cycling environments comfortable. However, this group is also more likely to rate *T_Local_Commercial* and *T_Major* –non-residential streets without bike lanes– as *somewhat comfortable. S*kipping ahead to the next section and observing Figure 1c we find that only 35% and 20% of respondents in this group rate these two videos as *very comfortable*, respectively. Hence, as a group, this segment cannot be qualified as being very comfortable. Finally, the least comfortable cyclists or non-cyclists dominate in the third cluster. This segment is most likely to find all descriptions uncomfortable. Residential streets, with and without sharrows, and trails are more likely to be rated as *somewhat uncomfortable* by the respondents in this group. Intent and cycling in the previous summer are significant variable-categories only for this group: respondents are more likely to strongly disagree with the intent statement (Table 2) and are more likely to not have cycled in the past summer; this group is therefore labeled Uncomfortable or Uninterested.

Overall, the characteristics of the three clusters in the 'Video Equivalent' segmentation are similar to the ones obtained in the 'Text Description' segmentation. The same observations about the roles of intent and cycling in the previous summer hold. The examination of comfort ratings, presented in Figure 1d also show the majority of respondents in the most comfortable group rate the least cyclist-friendly videos (*V_Local_Commercial* and *V_Major*) as *very comfortable*. These contrasts between the two segmentations prompted a change in nomenclature for the most comfortable group; they are named Very Comfortable Cyclists in the 'Video Equivalent' segmentation.

4.2 Demographic and Cycling Characteristics Comparison

The CCA results provide some insight as to the type of respondent that falls within each of the groups. However, to gain a better understanding of the characteristics of each group, we review demographic information, and analyze cycling characteristics and comfort ratings in more detail.

Table 6 shows several demographic characteristics for each of the Four Types of Cyclists, and the three cyclist types defined in the two empirical segmentations. Some notable observations include the gender disparity, particularly between the No Way No How and Strong and Fearless groups in the original typology, with men more represented in the Strong and Fearless and women, in the No Way No How. This disparity holds in the empirical segmentations, particularly between the Uncomfortable or Uninterested and (Very) Comfortable Cyclists, although to a lesser extent. There is also an age disparity, with older adults more present in the No Way No How group and, again, to a lesser extent in the Uncomfortable or Uninterested groups.

Variations in income and educational achievement between groups are more pronounced in the Four Types of Cyclists typology than in the empirical segmentations, and particularly the 'Video Equivalent' segmentation. Indeed, the proportion of higher income earners varies 19.9 percentage points from a low of 32.6% for No Way No How to a high of 52.5% for Enthused and Confident. On the other hand, the variation is only 9.7 percentage points between the Uncomfortable or

Uninterested (40.8%) and Very Comfortable Cyclists (50.5%). Similarly, the proportion of respondents who obtain a university degree varies by 23.9% between groups in the Four Types of Cyclists and only by 8.4% in the 'Video Equivalent' segmentation.

The immersive nature of the videos may explain some of the demographic differences noted by limiting the part imagination may play in assessing comfort, and potentially limiting the effect of varying reading proficiency between respondents.

Table 7 shows summary statistics regarding the intent and purpose of cycling as well as whether the respondent cycled in the last summer and when they last cycled. By definition, respondents in the No Way No How group cannot have cycled in the last summer and are likely to have been reclassified in the Interested but Concerned group if they agree with the intent statement, hence the low or null percentages observed for these variables. In contrast, a reasonable proportion (~ 40%) of Uncomfortable or Uninterested in both segmentations are interested in cycling more often, a reflection of the composition of this group, which includes respondents who are uncomfortable without necessarily being uninterested.

As expected, the Uncomfortable or Uninterested have the lowest proportion of respondents who agree with the intent statement within the empirical typologies. The highest proportion is found in the Cautious Majority. The comparatively lower proportion of (Very) Comfortable Cyclists who agree with the statement could be attributed to the already higher cycling frequency observed in this group (Figure 2); these respondents may be content with their current cycling levels. A similar trend for intent is found in the Four Types of Cyclists, where a lower percentage of Strong and Fearless agree they would like to cycle more often than Interested but Concerned or Enthused and Confident. However, a higher (and satisfying) cycling frequency is unlikely to explain the discrepancy between groups: 2.4% of Strong and Fearless have never cycled in their life, and less than 80% cycled in the previous year. In fact, proportionally, more Interested but Concerned cycled in the last year (88.3%) than Enthused and Confident (86.5%) or Strong and Fearless (79.2%). The proportion of respondents in each group that cycled in the last year follows a much more intuitive progression in the empirical segmentations, where the proportion increases with increasing comfort.

Table 6 Demographic Characteristics for Four Types of Cyclists and both Empirical Segmentations (Percent in Category)

	Four Types of Cyclists				**'Text Description' Segmentation**			**'Video Equivalent' Segmentation**		
Variable	*No Way No How*	*Interested but Concerned*	*Enthused and Confident*	*Strong and Fearless*	*Uncomfort-able or Uninterested*	*Cautious Majority*	*Comfortable Cyclists*	*Uncomfort-able or Uninterested*	*Cautious Majority*	*Very Comfortable Cyclists*
Gender										
Female	56.3[a]	45.7[b]	39.6[c]	25.6[d]	49.1[a]	48.1[a]	36.9[b]	52.6[a]	46.8[b]	31.3[c]
Male	39.5[a]	51.7[b]	56.7[b]	66.7[c]	45.1[a]	49.4[a]	59.7[b]	43.2[a]	50.7[b]	64.1[c]
Other°*	4.1	2.6	3.7	7.7	5.9[a]	2.5[b]	3.5[ab]	4.2[ab]	2.5[a]	4.6[b]
Age (Years)†										
15 – 24	2.5[a]	4.3[a]	4.5[a]	5.4[a]	3.3[a]	3.9[a]	5.0[a]	2.8[a]	4.5[a]	4.3[a]
25 – 44	25.1[a]	52.6[b]	46.3[c]	29.8[a]	39.2[a]	49.4[b]	45.2[c]	40.0[a]	51.3[b]	38.3[a]
45 – 64	41.4[a]	33.1[b]	39.6[a]	51.8[c]	35.9[a]	34.6[a]	39.2[a]	37.2[a]	32.8[b]	47.2[c]
65 – 98	29.4[a]	9.4[b]	9.3[b]	10.7[b]	19.7[a]	11.4[b]	10.0[b]	18.8[a]	10.8[b]	9.5[b]
Income ($ CAD)										
< 50,000	12.0[a]	8.8[a]	12.4[a]	12.5[a]	10.6[a]	9.3[a]	10.8[a]	10.0[a]	9.3[a]	11.7[a]
50,000 to 99,999	29.7[a]	27.3[a]	20.5[b]	17.3[b]	25.1[ab]	28.1[a]	22.4[b]	25.6[ab]	27.6[a]	21.9[b]
100,000 or more	32.6[a]	49.2[b]	52.5[b]	50.6[b]	40.1[a]	47.1[b]	51.9[c]	40.8[a]	48.6[b]	50.5[b]
Prefer not to answer	25.7[a]	14.7[b]	14.6[b]	19.6[ab]	24.2[a]	15.4[b]	14.8[b]	23.6[a]	14.4[b]	15.9[b]
Education†										
High school or less	16.3[a]	9.0[b]	9.8[b]	11.9[ab]	10.8[a]	10.0[a]	10.5[a]	11.7[a]	9.5[a]	11.5[a]
Technical school	33.6[a]	19.5[b]	21.6[b]	38.7[a]	31.5[a]	21.0[b]	22.2[b]	26.5[a]	21.2[b]	23.9[ab]
University degree	49.4[a]	70.9[b]	67.4[b]	47.0[a]	56.6[a]	68.4[b]	66.6[b]	60.5[a]	68.9[b]	63.3[a]

Note: All variables are statistically significantly different (Chi-square, $p < 0.05$).
[a, b, c,] Pairwise significant differences ($p < 0.05$). Cyclist types sharing the same letter for a given variable and within a typology are not statistically different.
* Low-count cells: Statistical significance must be considered with caution.
° Includes "Prefer not to answer" and "Neither describes me."
† *Prefer not to answer* category is negligible.

Table 7 Cycling Characteristics for Four Types of Cyclists and both Empirical Segmentations (Percent in Category)^

	Four Types of Cyclists				'Text Description' Segmentation			'Video Equivalent' Segmentation		
Variable	*No Way No How*	*Interested but Concerned*	*Enthused and Confident*	*Strong and Fearless*	*Uncomfort-able or Uninterested*	*Cautious Majority*	Comfortable Cyclists	*Uncomfort-able or Uninterested*	*Cautious Majority*	Very Comfortable Cyclists
Intent										
Somewhat or strongly agree	0.5[a]	79.0[b]	68.0[c]	35.1[d]	37.8[a]	70.1[b]	66.0[c]	43.4[a]	73.1[b]	57.3[c]
Purpose										
Recreation	5.5[a]	78.2[b]	77.2[b]	64.9[c]	44.1[a]	68.4[b]	77.8[c]	42.6[a]	73.4[b]	73.6[b]
Fitness	1.6[a]	46.9[b]	57.6[c]	45.8[b]	22.3[a]	40.2[b]	56.3[c]	20.6[a]	44.6[b]	56.5[c]
Utility	1.4[a]	55.9[b]	57.3[b]	37.5[c]	20.4[a]	48.8[b]	59.4[c]	21.2[a]	54.5[b]	52.2[b]
Commute	1.1[a]	46.8[b]	48.9[b]	24.4[c]	16.9[a]	41.8[b]	46.5[c]	16.8[a]	45.7[b]	42.7[b]
Biked in Summer 18										
Yes	0.0[a]	86.2[b]	83.4[b]	77.4[c]	48.4[a]	73.7[b]	87.3[c]	45.8[a]	79.9[b]	82.7[b]
Last Biked										
*Never**	4.8 [a]	0.4 [b]	1.1 [bc]	2.4 [ac]	3.3[a]	1.0[b]	0.6[b]	2.8[a]	0.8[b]	1.0[ab]
*In my childhood**	20.0 [a]	1.3 [b]	2.0 [b]	0.6 [b]	10.8[a]	3.6[b]	0.9[c]	10.5[a]	2.6[b]	1.2[b]
Several years ago	53.3 [a]	6.0 [b]	8.1 [b]	14.9 [c]	26.3[a]	12.4[b]	7.8[c]	28.3[a]	9.0[b]	11.3[b]
1-2 years ago	15.2 [a]	4.0 [b]	2.2 [b]	3.0 [b]	7.7[a]	6.3[a]	1.5[b]	8.3[a]	5.1[b]	2.3[c]
Within the last year	6.7 [a]	88.3 [b]	86.5 [bc]	79.2 [c]	51.9[a]	76.8[b]	89.2[c]	50.2[a]	82.5[b]	84.3[b]

Note: All variables are statistically significantly different (Chi-square, $p < 0.05$).
[a, b, c,] Pairwise significant differences ($p < 0.05$). Cyclist types sharing the same letter for a given variable and within a typology are not statistically different.
* Low-count cells: Statistical significance must be considered with caution.
^ See Table 2 for all variable descriptions.

Figure 1 shows comfort ratings for the descriptions and videos for the three typologies. Panels (a) and (b) show results of both text and video ratings by groups in the Four Types of Cyclists typology with the intent of providing some insight into the differences in rating between text-based and video-based input. Overall, videos are assessed in a similar manner to their equivalent text descriptions, with the exception of *V_Local_Commercial* and *V_Major* which are generally rated more favorably than their text counterparts. As the Four Types of Cyclists are classified using the text descriptions, only panel (a), along with panel (c) for the ‘Text Description’ segmentation and panel (d) for the ‘Video Equivalent’ segmentation are considered for the remainder of the comparison.

As expected, Figure 1 shows that mixed traffic cycling environments on non-residential streets (*T_Local_Commercial*, *T_Major*, and their video equivalents) are least comfortable for all groups, except the Strong and Fearless, who, by definition, must rate these two descriptions as *very comfortable* to be part of the group. Generally, separated trails (*T_Path, V_Path*) are considered *very comfortable* by more respondents than the protected bike lanes (*T_Major_PBL, V_Major_PBL*). Note that there is an interesting discrepancy in the Strong and Fearless group, who proportionally have more respondents who rate the above-mentioned separated paths and protected bike lanes as *somewhat uncomfortable* or *very uncomfortable* compared to the Interested but Concerned and Enthused and Confident. We hypothesize that is a reflection of the Strong and Fearless’ likely preference for speed and directness. Members of this group likely engage in vehicular cycling; the windiness (and thus, interrupted sightlines) of paths and conflicts with other users, particularly pedestrians, as well as frequent stops or crowdedness of protected lanes may reduce their perception of comfort on these infrastructures. This discrepancy is not reflected in the empirical segmentations. We hypothesize this is because the highest comfort group in both segmentations includes more cyclists than the stereotypical Strong and Fearless; the particularities of these Strong and Fearless-type cyclists were not sufficiently different from the other high-comfort respondents to lead to a distinctive and stable cluster when applying CCA. Nonetheless, a spectrum of comfort levels is expected to exist within each cluster and the particular preferences of vehicular or Strong and Fearless-type cyclists are therefore likely to be masked by characteristics of the rest of the (Very) Comfortable Cyclist groups.

In addition, the empirical segmentations yield more homogeneous groups when it comes to the comfort ratings compared to the classic Four Types of Cyclists. This is particularly noticeable for the Uncomfortable or Uninterested groups; as compared to the No Way No How, the empirical clusters for both segmentations group cyclists who are mostly *very uncomfortable* on a majority of facilities, which is not the case in the Four Types classification. This categorization of cyclists who are generally very uncomfortable in the Uncomfortable or Uninterested group highlights that the unifying factor for this cyclist type is low comfort, while being uninterested is a salient but optional characteristic. Recall from Table 7 that about 40% of the group are in fact interested in cycling more often.

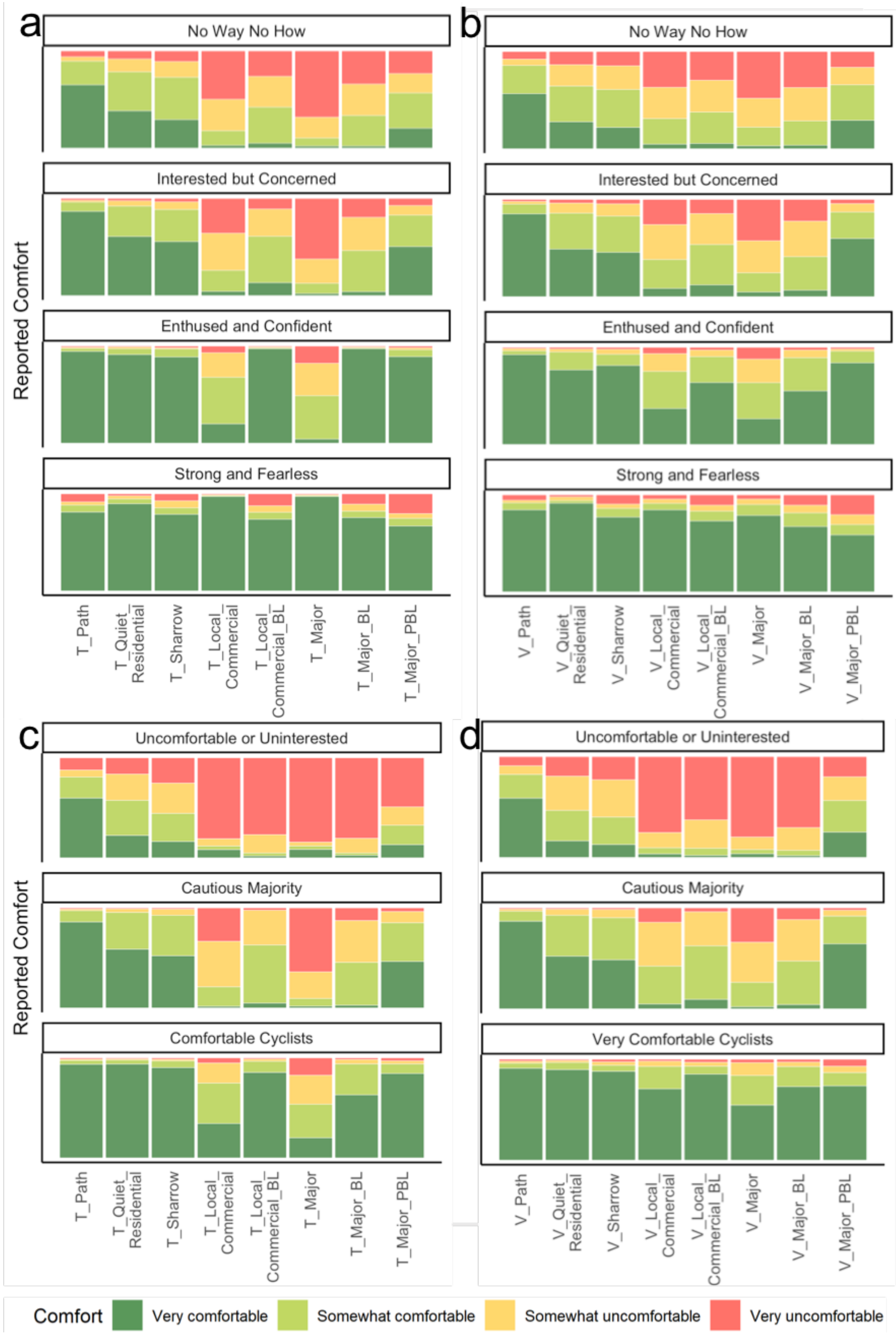


Figure 1 Comfort Ratings of Four Types of Cyclists for (a) Facility Descriptions and (b) Videos, of (c) Descriptions for 'Text Description' Segmentation, and of (d) Videos for 'Video Equivalent' Segmentation

Figure 2 shows the cycling frequency by season for each cyclist type in the Four Types of Cyclists and the two empirical segmentations. On a season basis, frequencies are statistically significantly different (Chi-square, $p < 0.05$) between groups. In line with observations from the original investigations in the Four Types of Cyclists (Dill and McNeil, 2013, 2016), the first panel of Figure 2 shows that higher comfort does not necessarily correlate to higher cycling frequency, excluding the No Way No How group. In fact, in all seasons but winter, the proportion of those who cycle very regularly (at least two to three times per week) is higher for both the Interested but Concerned and the Enthused and Confident as compared to the Strong and Fearless. This known critique of the typology is reflected in our data.

One hypothesis to explain this comparatively low cycling frequency is that some cyclists who fall in the Strong and Fearless category are likely to be weekend road cyclists who use highways and rural roads, and thus have a high tolerance for roadways without cycling-specific infrastructure while also not cycling at a very high frequency. In contrast, for the two empirical segmentations, Figure 2 shows that frequency increases with greater comfort. This does not mean that Uncomfortable or Uninterested cyclists never cycle; on the contrary, we can find daily cyclists in all categories. However, the proportion of daily or frequent cyclists grows, as expected, with increased cycling comfort. We can hypothesize that the uncomfortable cyclists who cycle frequently may do so recreationally, or may live in a neighborhood with an adequate supply of infrastructure that is considered highly comfortable across the entire population, enabling them to reach their main destinations. As the survey instrument did not record cycling frequency for each cycling purpose individually, the possible explanations put forward remain hypotheses that should be confirmed in a future survey.

Finally, another interesting observation from Figure 2 is the winter cycling rate. While cycling frequency largely drops in winter for all cyclist types across all three typologies, relatively high cycling frequency (once a week or more) follows a pronounced progression with comfort level. Based on the ‘Video Equivalent’ segmentation, 7.4% (n = 46) of Uncomfortable or Uninterested, 15.9% (n = 330) of the Cautious Majority, and 21.9% (n = 113) of Very Comfortable Cyclists continue to cycle at least once a week in winter. However, as with the general cycling frequency trend, the progression is not observed in the Four Types of Cyclists typology: 0.9% (n = 4) of No Way No How, 17.6% (n = 395) of Interested but Concerned, 17.7% (n = 63) of Enthused and Confident, and 16.1% (n = 27) of Strong and Fearless cycled at least once a week during the previous winter. Cycling in winter requires a few additional skills to adjust to slippery, snowy, and dark conditions. Higher comfort as measured through our survey instrument may be correlated with higher comfort in these winter conditions, which would explain the progression observed in the empirical segmentations.

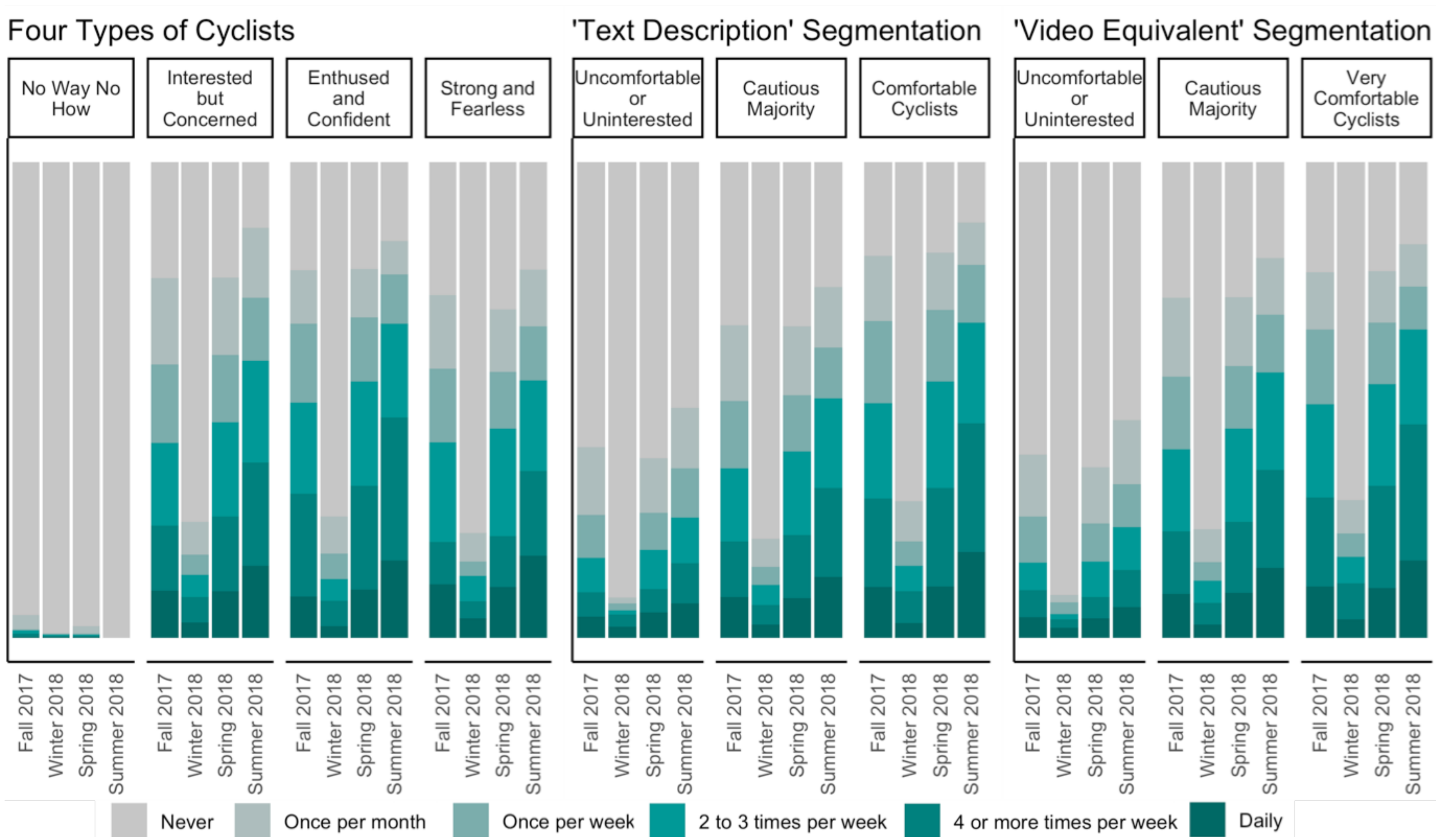


Figure 2 Proportion of the Stated Cycling Frequency per Season in the Previous Year for Each Cyclist Type in the Three Typologies

4.3 Reclassification of Respondents Between the Four Types of Cyclists and the Two Empirical Segmentations

The two typologies obtained using CCA differ in several ways from the Four Types of Cyclists. The most obvious distinction is the presence of three classes rather than four. This finding is consistent with results of our previous work, where a factor analysis on facility descriptions resulted in three groups being defined (Cabral et al., 2018). A cross-classification of the membership in the Four Types of Cyclists against the two empirical segmentations (Table 8) shows two important redistributions of the class membership. First, only 34.9%/49.7% ('Text Description'/'Video Equivalent', respectively) of No Way No How are classified as Uncomfortable or Uninterested (row percentages, Table 8). Second, the Uncomfortable or Uninterested cluster is composed of 35.7%/35.0% No Way No How and 58.5%/60.8% Interested but Concerned (column percentages, Table 8). The redistribution is also present for the next level of cyclists, although in a less pronounced way: 75.8%/76.6% of the Interested but Concerned are reclassified as Cautious Majority (row percentages), while the Cautious Majority cluster is composed at 86.1%/83.0% of respondents categorized as Interested but Concerned (column percentages). Most of the Enthused and Confident and Strong and Fearless fall in the third cluster, although to a lesser extent in the 'Video Equivalent' segmentation.

Table 8 Cross-classification between Four Types of Cyclists and Empirical Segmentations
(cell percentage, row percentage, *column percentage*)

	'Text Description' Segmentation			'Video Equivalent' Segmentation		
	Uncomfortable or Uninterested	*Cautious Majority*	*Comfortable Cyclists*	*Uncomfortable or Uninterested*	*Cautious Majority*	*Very Comfortable Cyclists*
No Way No How	4.7 34.9 *35.7*	8.5 62.8 *13.8*	0.3 2.3 *1.2*	6.7 49.7 *35.0*	6.4 47.1 *9.9*	0.4 3.2 *2.7*
Interested but Concerned	7.8 11.1 *58.5*	53.1 75.8 *86.1*	9.2 13.2 *36.9*	11.7 16.7 *60.8*	53.7 76.6 *83.0*	4.7 6.7 *29.3*
Enthused and Confident	<0.1 *0.3* *0.2*	<0.1 *0.3* *0.1*	11.0 99.4 *44.1*	0.3 2.8 *1.6*	4.3 39.0 *6.7*	6.5 58.1 *40.2*
Strong and Fearless	0.7 14.3 *5.6*	<0.1 *0.6* 0.1	4.5 85.1 *17.8*	0.5 9.5 *2.6*	0.3 5.4 *0.4*	4.5 85.1 *27.8*

Table 8 contains some surprising reclassifications: a small number of respondents considered to be No Way No How are reclassified as Comfortable Cyclists or Very Comfortable Cyclists in the empirical segmentations and, conversely, some Strong and Fearless are categorized as Uncomfortable or Uninterested. A closer look at the first case type indicates the respondents did not rate both *T_Local_Commercial* and its bike lane variant as *very comfortable,* and neither *T_Major* and its bike lane variants, which means they were automatically considered either No Way No How or Interested but Concerned as per Dill and McNeil's methodology. As none had cycled in the last two years and most disagree they would like to cycle more often, they were classified as No Way No How. The ratings for the videos equivalent to those four descriptions were much more positive, which explains why the 'Video Equivalent' segmentation classified those respondents as Very Comfortable Cyclists. As for the 'Text Description' segmentation the CCA likely classified those cyclists as comfortable as they all rated at least one, and sometimes two of the descriptions as *very comfortable*. It could very well be argued that these respondents belong more intuitively in the Uncomfortable or Uninterested group since they do not cycle and have no intention of cycling in the future.

On the other hand, the second set of cases contains respondents who rated both *T_Local_Commercial* and *T_Major* as *very comfortable*, automatically classifying them as Strong and Fearless. Two scenarios emerge when looking at the specific responses for those respondents. For one group, their ratings of the equivalent videos were on the uncomfortable side, as were many other video ratings. This subgroup also strongly disagreed they would like to cycle more often. Arguably, they are best classified as Uncomfortable or Uninterested rather than as Strong and Fearless. The other group had a tendency to rate *T_Path* and *T_Major_PBL* as uncomfortable and had response patterns that suggested a preference for vehicular cycling, although not all in the group had cycled in the past year. Some in this sub-group seem to fit the description of Uninterested, while others would be better classified as (Very) Comfortable Cyclists. These discrepancies point to the difficulties inherent in creating summary categories to describe a wide variety of different cyclists and non-cyclists in one unified framework. No typology is able to fully capture the variety of cyclist comfort, intent, and cycling frequency. However, our methodology defines within-group comfort rating patterns that are more homogeneous than that of the Four Types of Cyclists (Figure 1), particularly in the 'Video Equivalent' segmentation.

Comparing the two empirical segmentations, there are some reclassifications, albeit not as important as with the Four Types classification. Compared to the 'Text Description' segmentation, the proportion of respondents in each cluster of the 'Video Equivalent' segmentation changes: the Cautious Majority cluster gains membership (from n = 1979 to n = 2075), and the Uncomfortable or Uninterested cluster (n = 618) becomes larger than the Very Comfortable Cyclists (n = 515). Consistent with the changes in proportions, there is a reclassification of some Comfortable Cyclists as Cautious Majority, and of some in the Cautious Majority group as Uncomfortable or Uninterested in the 'Video Equivalent' segmentation. The most comfortable cluster shrinks in size from 25% to 16% of respondents. These reclassifications result in more homogeneous groups in the 'Video Equivalent' segmentation compared to the 'Text Description' segmentation, particularly for the (Very) Comfortable Cyclists group.

5 CONCLUSIONS

Our work uncovers some new and important empirical findings about the Four Types of Cyclists typology. First, and most critically, we find the rule-based method by which survey respondents are classified into the four types yields quite heterogeneous groups, particularly with respect to perceived comfort. We also find that intent plays a fairly minor role in defining cyclist type. This is a relatively surprizing result given intent is a known determinant of the decision to cycle, often mediated by habit (Danner et al., 2008; de Bruijn et al., 2009). One possible explanation for the more minor role of intent in defining cyclist types is the clustering method itself: only one of the ten variables included was intent, whereas eight of the variables were related to infrastructure comfort. Nonetheless, intent is not entirely absent of cyclist type definition; lack of intent to cycle is a defining factor for the Uncomfortable or Uninterested.

While we highlight many differences between the classic Four Types and our empirical segmentations, there are several similarities between the typologies. Notably, the observed gender distributions, where more men are present in the makeup of the more comfortable cyclist groups, and more women in the less comfortable or non-cycling groups are reflected in all typologies. This trend was to be expected and is in line with the literature on the subject (Heinen et al., 2010; Sener et al., 2009)

Our empirical segmentation leads to three cyclist types instead of four: 1) Uncomfortable or Uninterested, 2) Cautious Majority, and 3) (Very) Comfortable Cyclists. Our analysis highlights some differences in the typology when derived from text descriptions versus videos. For example, the highest comfort group in the text-based segmentation is not as distinctly more comfortable than the Cautious Majority group compared to the video-based segmentation, which has a clearly defined very comfortable cyclist group. For practitioners wishing to reproduce this work in their own locale, we suggest using videos is likely to yield slightly more accurate results due to the advantages of videos previously stated, such as reducing barriers for respondents with lower reading proficiency and presenting all the same contextual cues for all respondents. However, the high level of similarity between the text-based segmentation and the video-based segmentation suggests that adequate results would still be obtained using only descriptions, particularly for high-level planning projects. Indeed, the most important characteristics of the typology, namely that it is composed of three cyclist types and the general characteristics of these cyclists are common for both segmentations.

This new three-group typology fulfills the other main goal of our research – to define cyclist types that can be used to inform policies regarding infrastructure choice. Key findings regarding comfort on different types of infrastructure will inform a reassessment of the LTS framework, which we explore in the next phase of research. First, Uncomfortable or Uninterested are comfortable on separated trails, but protected bike lanes are not perceived to offer a sufficient level of separation to make them feel at ease. Second, the Cautious Majority forms a large group where some show preference for separated facilities (protected bike lanes), while others are comfortable on calm residential streets. Painted bike lanes and contra-flow lanes are not perceived as comfortable by this group. The provision of a strong network of protected lanes with feeder residential streets would likely be most suitable for these (potential) cyclists. Third, most Very Comfortable Cyclists are not fearless cyclists: some dedicated infrastructure, including painted bike lanes on major roads, can help increase comfort for this group.

Although it might seem unintuitive to have a typology devoid of a dedicated non-cyclist category similar to the No Way No How, we do not believe this to be problematic. Indeed, as we noted in our introduction, we are particularly interested in a typology that can be used as a basis for a reassessment of the LTS framework, where cyclist types have relatively homogeneous infrastructure preferences. The clustering results suggest those unwilling to cycle are also more likely to be uncomfortable in many cycling environments, and thus have similar infrastructure preferences. While we did not aim to quantify potential cycling demand, jurisdictions where this is a question of importance may find this feature of our typology to be limiting. One may consider including further survey questions regarding physical ability and likelihood of cycling in the future to obtain this information without including a dedicated non-cycling group in the typology.

In addition, as pointed out by Damant-Sirois et al. (2014), the Four Types of Cyclists typology is mostly useful to formulate policy recommendations regarding infrastructure. To encourage the initiation of cycling among those who are very uncomfortable or unwilling to cycle, in addition to providing adequate infrastructure, policies would need to target other aspects of the choice to cycle (cultural biases, education, financial incentives, land use densities, etc.). In terms of infrastructure, the only comfortable environment for this very uncomfortable or unwilling group appears to be completely segregated cycling or multi-use (walking and cycling) trails. Policies aimed at building or upgrading trails should therefore be pursued to accommodate respondents in this group.

Our work has several limitations, the first of which is the non-random sampling method used to obtain survey responses. Given the large sample size, we feel the typology itself is likely fairly representative of the total population. However, the proportion of respondents in each type are likely not reflective of the population as a whole. In particular, we expect the Uncomfortable or Uninterested to be underrepresented in our sample. The survey was also administered online and in English only, which limits *de facto* the ability of a certain number of Edmontonians to answer, in particular lower-income groups, non-English speakers, visually impaired Edmontonians, and those who do not have the required computer or reading literacy. In addition to repeating this survey with a random sample in Edmonton, it would be beneficial to reproduce the survey in other locales to verify if the typology is transferable, or if it reflects particularities of the cycling infrastructure available in Edmonton and of its cycling and non-cycling population.

Another limitation is the introduction of some variations to the questionnaire used by Dill and McNeil (2013), including changes in wording for the facility descriptions to reflect a local Canadian context and a reduction in the number of statements. Our criteria to define current cyclists also differed: rather than cycling at least once in the last month, our method included those who cycled at least once a month in the previous season (summer 2018) as current cyclists. These differences may in part explain some of the variations observed in our work, both in the evaluation of the Four Types of Cyclists, and in the subsequent empirical segmentations.

Further, our survey did not include questions regarding ability to ride. It is expected that those unable to ride because of physical ability or lack of learning opportunity will have indicated that they never ride a bicycle. However, it is unclear how riding ability would influence comfort ratings. It is likely that comfort perception would be influenced by the particular circumstances of each respondent. For example, a respondent with previous bicycling experience who can no longer ride may have used their memory to answer the comfort questions, leading to varied response patterns. Respondents who cannot currently ride, but who expect to ride in the future (when a temporary disability is removed or when they learn to ride, for example) would likely try to project themselves in the future and imagine how comfortable they would feel. A future iteration of this survey should include questions regarding ability to ride and evaluate potential impacts on the typology.

We are currently evaluating the correspondence between stated comfort level and actual route choice, as it is well known that stated responses and actual preferences can differ significantly (Wardman, 1988). This work will make use of an optional survey module, where one hundred respondents who cycled at least once in the last year were asked to describe their most common utility or commute cycling route. Results from this work can also contribute to redefining the LTS framework. Further exploratory segmentations using the full set of 16 video clips and other variables such as cycling frequency and purpose are also part of our ongoing work.

ACKNOWLEDGMENTS

The authors would like to thank the City of Edmonton for their collaboration and support for this project. We would also like to acknowledge research group members who helped with data collection and survey testing: Matthew Woo, Wenxin Wang, Bryan Tran, and Sabrena Jahan Ohi.

Funding: This work was supported by the City of Edmonton [project number RES0035313]. Laura Cabral was supported by the Natural Sciences and Engineering Research Council of Canada and the Government of Alberta.

REFERENCES

Ajzen, I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 50, 179-211.

Blanc, B., Figliozzi, M. (2016) Modeling the impacts of facility type, trip characteristics, and trip stressors on cyclists' comfort levels utilizing crowdsourced data. *Transportation Research Record: Journal of the Transportation Research Board* 2587, 100-108.

Cabral, L. (2019) Analyzing network connectivity in light of cyclist comfort: An empirical reappraisal of the four types of cyclists and of the level of traffic stress framework. Master's Thesis, University of Alberta.

Cabral, L., Kim, A.M., Parkins, J.R. (2018) Bicycle ridership and intention in a northern, low-cycling city. *Travel Behaviour and Society* 13, 165-173.

Cabral, L., Kim, A.M., Shirgaokar, M. (2019) Low-stress bicycling connectivity: Assessment of the network build-out in Edmonton, Canada. *Case Studies on Transport Policy* 7, 230-238.

Chen, C., Anderson, J.C., Wang, H., Wang, Y., Vogt, R., Hernandez, S. (2017) How bicycle level of traffic stress correlate with reported cyclist accidents injury severities: A geospatial and mixed logit analysis. *Accident Analysis & Prevention* 108, 234-244.

Damant-Sirois, G., El-Geneidy, A.M. (2015) Who cycles more? Determining cycling frequency through a segmentation approach in Montreal, Canada. *Transportation Research Part A: Policy and Practice* 77, 113-125.

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

Danner, U.N., Aarts, H., de Vries, N.K. (2008) Habit vs. intention in the prediction of future behaviour: the role of frequency, context stability and mental accessibility of past behaviour. *British Journal of Social Psychology* 47, 245-265.

de Bruijn, G.J., Kremers, S.P., Singh, A., van den Putte, B., van Mechelen, W. (2009) Adult active transportation: adding habit strength to the theory of planned behavior. *American Journal of Preventive Medicine* 36, 189-194.

Dill, J., McNeil, N. (2013) Four types of cyclists? Examining a typology to better understand bicycling behavior and potential. *Transportation Research Record: Journal of the Transportation Research Board* 2387, 129-138.

Dill, J., McNeil, N. (2016) Revisiting the four types of cyclists. *Transportation Research Record: Journal of the Transportation Research Board* 2587, 90-99.

Dill, J., Voros, K. (2007) Factors affecting bicycling demand: Initial survey findings from the Portland, Oregon, region. *Transportation Research Record: Journal of the Transportation Research Board* 2031, 9-17.

Dolnicar, S., Grün, B. (2008) Challenging "factor–cluster segmentation". *Journal of Travel Research* 47, 63-71.

Eriksson, L., Forward, S.E. (2011) Is the intention to travel in a pro-environmental manner and the intention to use the car determined by different factors? *Transportation Research Part D: Transport and Environment* 16, 372-376.

Félix, R., Moura, F., Clifton, K.J. (2017) Typologies of urban cyclists: Review of market segmentation methods for planning practice. *Transportation Research Record: Journal of the Transportation Research Board* 2662, 125-133.

Gatersleben, B., Haddad, H. (2010) Who is the typical bicyclist? *Transportation Research Part F: Traffic Psychology and Behaviour* 13, 41-48.

Geller, R. (2006) Four Types of Cyclists.

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.

Harkey, D.L., Reinfurt, D.W., Knuiman, M. (1998) Development of the Bicycle Compatibility Index. *Transportation Research Record: Journal of the Transportation Research Board* 1636.

Heinen, E. (2016) Identity and travel behaviour: A cross-sectional study on commute mode choice and intention to change. *Transportation Research Part F: Traffic Psychology and Behaviour* 43, 238-253.

Heinen, E., van Wee, B., Maat, K. (2010) Commuting by bicycle: An overview of the literature. *Transport Reviews* 30, 59-96.

Hennig, C. (2007) Cluster-wise assessment of cluster stability. *Computational Statistics & Data Analysis* 52, 258-271.

Jensen, S. (2007) Pedestrian and bicyclist level of service on roadway segments. *Transportation Research Record: Journal of the Transportation Research Board* 2031, 43-51.

Kent, M., Karner, A. (2018) Prioritizing low-stress and equitable bicycle infrastructure using a novel accessibility measure. *Transportation Research Board 97th Annual Meeting*, Washington DC.

Kroesen, M., Handy, S. (2014) The relation between bicycle commuting and non-work cycling: results from a mobility panel. *Transportation* 41, 507-527.

Landis, B.W., Vattikuti, V.R., Brannick, M.T. (1997) Real-time human perceptions: Toward a Bicycle Level of Service. *Transportation Research Record: Journal of the Transportation Research Board* 1578, 119-126.

Larsen, J., El-Geneidy, A. (2011) A travel behavior analysis of urban cycling facilities in Montréal, Canada. *Transportation Research Part D: Transport and Environment* 16, 172-177.

Lehtonen, E., Havia, V., Kovanen, A., Leminen, M., Saure, E. (2016) Evaluating bicyclists' risk perception using video clips: Comparison of frequent and infrequent city cyclists. *Transportation Research Part F: Traffic Psychology and Behaviour* 41, 195-203.

Li, Z., Wang, W., Liu, P., Ragland, D.R. (2012) Physical environments influencing bicyclists' perception of comfort on separated and on-street bicycle facilities. *Transportation Research Part D: Transport and Environment* 17, 256-261.

Lois, D., Moriano, J.A., Rondinella, G. (2015) Cycle commuting intention: A model based on theory of planned behaviour and social identity. *Transportation Research Part F: Traffic Psychology and Behaviour* 32, 101-113.

Markos, A., D'Enza, A.I., Velden, M.v.d. (2018) Beyond tandem analysis: Joint dimension reduction and clustering in R. *Journal of Statistical Software*. (In Press).

Markos, A., Iodice D'Enza, A., van de Velden, M. (2019) clustrd: Methods for joint dimension reduction and clustering. R package version 1.3.1.

Mekuria, M.C., Furth, P.G., Nixon, H. (2012) Low-stress bicycling and network connectivity. Mineta Transportation Institute Publications.

Moran, S.K., Tsay, W., Lawrence, S., Krykewycz, G.R. (2018) Lowering bicycle stress one link at a time: Where should we invest in infrastructure? *Transportation Research Record: Journal of the Transportation Research Board* 2627, 33-41.

Papageorgiou, G., Markos, A., Zarkadis, N. (2016) Understanding the atom and relevant misconceptions: Students' profiles in relation to three cognitive variables. *Science Education International* 27, 464-488.

Parkin, J., Wardman, M., Page, M. (2007) Models of perceived cycling risk and route acceptability. *Accident Analysis & Prevention* 39, 364-371.

Rousseeuw, P.J. (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* 20, 53-65.

Semler, C., Sanders, M., Buck, D., Dock, S., Cesme, B., Wang, S. (2018) The keys to connectivity: The District of Columbia's innovative approach to unlocking low stress bicycle networks. *Transportation Research Board 97th Annual Meeting*, Washington DC, p. 16.

Sener, I., Eluru, N., Bhat, C. (2009) Who are bicyclists? Why and how much are they bicycling? *Transportation Research Record: Journal of the Transportation Research Board* 2134, 63-72.

Shaw, R.A., Crane, J., Pearce, N., Burgess, C.D., Bremner, P., Woodman, K., Beasley, R. (1992) Comparison of a video questionnaire with the IUATLD written questionnaire for measuring asthma prevalence. *Clinical and Experimental Allergy* 22, 561-568.

Sheeran, P. (2002) Intention—behavior relations: A conceptual and empirical review. *European Review of Social Psychology* 12, 1-36.

Sleed, M., Durrheim, K., Kriel, A., Solomon, V., Baxter, V. (2002) The effectiveness of the vignette methodology: A comparison of written and video vignettes in eliciting responses about date rape. *South African Journal of Psychology* 32, 21-28.

Smith, T.W., Sokolowski, J. (2008) *Using Audio-Visuals in Surveys*. NORC/University of Chicago.

Statistics Canada (2017) Census Profile, 2016 Census.

Steinley, D. (2004) Properties of the Hubert-Arabie adjusted Rand index. *Psychological Methods* 9, 386-396.

Stinson, M.A., Bhat, C.R. (2004) Frequency of bicycle commuting: Internet-based survey analysis. *Transportation Research Record: Journal of the Transportation Research Board* 1, 122-130.

Stinson, M.A., Bhat, C.R. (2005) A comparison of the route preferences of experienced and inexperienced bicycle commuters *Transportation Research Board 84th Annual Meeting*, Washington, D.C.

van de Velden, M., D'Enza, A.I., Palumbo, F. (2017) Cluster Correspondence Analysis. *Psychometrika* 82, 158-185.

Veillette, M.-P., Grisé, E., El-Geneidy, A. (2019) Does one bicycle facility type fit all? Evaluating the stated usage of different types of bicycle facilities among cyclists in Quebec City, Canada. *Transportation Research Record: Journal of the Transportation Research Board*.

Wardman, M. (1988) A comparison of revealed preference and stated preference models of travel behaviour. *Journal of Transport Economics and Policy* 22, 71-91.

Wilkinson, W.C.A., Clarke, B., Epperson, B., Knoblauch, R. (1994) Selecting roadway design treatments to accommodate bicycles. Federal Highway Administration, U.S. Department of Transportation.

Winters, M., Davidson, G., Kao, D., Teschke, K. (2011) Motivators and deterrents of bicycling: Comparing influences on decisions to ride. *Transportation* 38, 153-168.