Evaluating Pedestrian Preference of Walkable Streets using a Stated Preference Survey

Written by: Sneha Adhikari

ABSTRACT

The research paper analyzes pedestrian perception and preferences of street-level attributes found in the City of Toronto. The purpose of the study is to uncover the existing gap of information on pedestrian experience found in urban design practice and prior studies that can shape the decision-making of streetscape design choices. The study uses a stated preference survey to assess choice on a set of hypothetical scenarios found within three street attributes: through lane, curb lane, and sidewalk. The data is estimated as a choice model using a mixed multinomial logit model to come to conclusions on the effects of specific elements found within the attributes on pedestrian perception. The results of the paper establish a utility model with all significant explanatory variables that quantifies pedestrian experience on the chosen set of scenarios.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Professor Matt Roorda and Dena Kasraian, for all their support, expert knowledge, and for trusting me with such a big project. This paper would not be where it is without their assistance and encouragements. I am also thankful for Prof. Roorda's research students who provided a lot of continual feedback throughout the progression of this year-long study. In addition, I want to thank the brilliant Jason Hawkins who helped me understand the concepts behind choice modelling upon which this paper relies on. The experience of working alongside such intelligent researchers within the University and getting to work on a project with real-world impact has been a unique experience to an undergraduate student like myself with an interest in pursuing transportation engineering in the future. The lessons I have learned on this year-long study has enriched my knowledge and will be something I will bring with me throughout my career.

Lastly, I want to extend my thanks to my parents who have always been a beacon of support and encouragement. I am also grateful to my friends in the civil engineering program for supporting me through the years so that I felt confident about taking up an opportunity like this one.

Contents

1.0 INTRODUCTION

1.1 Preface

The following document is a part of the *iCity* project (iCity: Urban Informatics for Sustainable Metropolitan Growth, funded by The Ontario Research Fund – Research Excellence Round 7), work package 2.1 'Evaluation of Complete Streets' (Project 2.1 Evaluation of Complete Streets, 2017). The original project aims to quantify Complete Streets and provide guidelines for their design. This is an interdisciplinary collaboration between Waterfront Toronto, University of Toronto, and visualization experts from OCAD University and industry partner, ESRI Canada.

1.2 Background

Assessment of urban streetscape designs are constantly evolving. Many cities have urban design guidelines, but they are rarely based on empirical evidence and tend to focus on vehicular traffic at street level. In the context of complete streets, for example, most existing policy and literature only touch on qualitative goals of complete streets, before-and-after effects of implementation, and design elements relating to level-of-service (LOS) attributes (Hui, Roorda, Saxe, Hess, & Miller, 2017). Some complete street policies include performance measures and make mention of context-sensitivity (National Complete Streets Coalition, 2016). However, these measurements do not include a quantification of pedestrian comfort or experience on streets (Hui, Roorda, Saxe, Hess, & Miller, 2017).

Many cities, including the City of Toronto, collect and update data on traffic volumes as well as pedestrian counts at specific intersections and corridors (City of Toronto, 2018). These counts provide an indication of pedestrian activity, but the missing component of behavior and psychology prevails. Moreover, such behavior-focused data is usually in the form of public opinion polls that do not have a robust empirical measure, is often specific to a region where a proposed project is taking place, and is intended to collect feedback on existing infrastructure. Additionally, there is little prior research done that demonstrate a trade-off analysis between various design alternatives that could be found within a street.

This research paper aims to address this lack of available information and research to quantify pedestrian experience within a streetscape. For the study, a stated preference (SP) survey is chosen as the best methodology as it allows for the testing of hypothetical scenarios with choice-based outcomes. The goal of this research is to quantify pedestrian preference through choice probabilities of street attributes presented to respondents.

2.0 LITERATURE REVIEW

SP research have been used to assess behavior and choice when a new set of hypothetical scenarios are to be introduced and are often also used to forecast decisions. In transportation, SP surveys have been used to assess respondent choice for different modal alternatives based on their attributes (i.e. travel time, cost, wait time) which gives information on the way choice is made (Hensher, 1994). For the context of this research, the survey design includes a ranking system where the respondents are asked to rank their choice among multiple streetscape scenarios. The following section includes examples of prior research using SP surveys that also study pedestrian preference – however, nothing currently exists that measures preference within the specific street attributes presented in this paper.

2.1 Pedestrian Preferences with Respect to Roundabouts

Prior studies measuring pedestrian experience using a SP survey exist with much of it focused on specific design considerations including crosswalks, and roundabouts. A joint research team from several Quebec universities, Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation, published a 2013 study “Pedestrian Preferences with Respect to Roundabouts - A Video-based Stated Preference Survey” (Perdomo, Rezaei, Patterson, Saunier, & Miranda-Moreno, 2014). The study used a web-based video simulation for a stated preference survey in the context of roundabout scenarios in Montreal with results presented as a panel mixed multinomial logit (MMNL) model. The findings show that vulnerable road users (VRUs) preferred roundabouts with crosswalk due to safety concerns and that they had a higher preference for lower speed and volume, as well as a high preference of flashing pedestrian crossing signs over no signs (Perdomo, Rezaei, Patterson, Saunier, & Miranda-Moreno, 2014). This study, specifically the methodology of using a web-based stated preference survey and analysis with a panel logit model, provides a good insight as a reference for this research.

2.2 Estimating preferences for different types of pedestrain crossing facilities

There have been numerous studies that look at pedestrian perception of crossing facilities. Researchers from University College London conducted a SP survey in three English cities where respondents were asked to indicate comfort with respect to different types of crossing facilities and then choose between walking for additional time to use a type of crossing facility or avoid crossing the road. Using mixed logit modelling, they found that participants, on average, were willing to walk an additional 2.4 and 5.3 min to use a signalized crossing and avoid using underpasses and footbridges (Anciaes & Jones, 2017). This study was specific to crossing facilities, but the application of stated preference-based research and use of mixed logit modelling provides a benchmark basis for the present study.

3.0 METHODOLOGY

3.1 Survey instrument and scenarios

The scope of the survey is narrowed to include attributes at street segment level for the purposes of recreational walking. The survey instrument is a combined effort by various parties in the iCity group. The instrument is web-based with questions that ask respondents to rate an existing street (revealed preference) but also re-rate hypothetical manipulated options (stated preference). For the revealed preference case, several photos of streets at Toronto waterfront and downtown are shown and respondents rate their preference of these photos individually. The larger portion of the survey includes the stated preference assessment. In order to simulate theoretical scenarios, ESRI's visualization team used CityEngine & Unity to create animations modelled after four chosen streets in Toronto. The final version of the survey include streets modelled after existing streets but with details like colours and building facade dulled out to ensure that respondents' selection remain unbiased without any influence from variables outside of the ones being measured (i.e. the type of building facade might affect respondents' idea of how liveable or safe the street might look which is outside the scope of what the study is trying to measure). Listed below is a full breakdown on the survey components:

- Five questions regarding respondent demographics: place of residence, age, gender, and familiarity with the Toronto Waterfront
- Four questions assessing revealed preference based on photographs of Toronto's existing streets
- Nine sets of questions assessing stated preference on hypothetical animated scenarios
- 14 questions on respondent demographics: employment status, occupation category, education, household characteristics, daily transportation and cycling habits, access to a driver's license and personal vehicle, and information on walking pace/difficulty and vision.

The stated preference portion of the survey present respondents with a choice of three street scenarios at a time with nine sets of total questions. Therefore, there are 27 different scenarios that can be chosen. Each scenario is animated and display different attribute levels that fall within three street attributes: through lane, curb lane and sidewalk. **Table 1** outlines all the attributes and their associated attribute levels.

Table 1: Survey attributes (through lane, curb lane, and sidewalk) and their associated levels being tested

	Through lane	**Curb lane**	**Sidewalk**
Attribute levels	car + car + car + car	two-way cycle path	trees + 1.6m pedestrian walkway
	car + car	one-way cycle path	trees + 3.2m pedestrian walkway
	car + transit + transit + car	on street parking + one-way cycle path	curbside outdoor dining and trees + 1.6m pedestrian walkway
	transit + car + car + transit	one-way cycle path + on street parking	1.6m pedestrian walkway + wall-side outdoor dining
	transit + transit	none	3.2m pedestrian walkway

These attribute levels are combined to form the scenarios found in the survey. **Figure 1** shows a snapshot of the online survey with some of the combined attributes levels. The scenarios are presented three at a time and respondents are asked to rank their first and second choice, as well as their confidence level with their ranking decision. As respondents are asked to make repeated choices within a specific period of time, a panel data analysis is necessary. For this study, a panel mixed multinomial logit modelling approach is selected to assess the choice probabilities, as explored in **Section 3.3** Multinomial Logit Model and Mixed Logit Model.

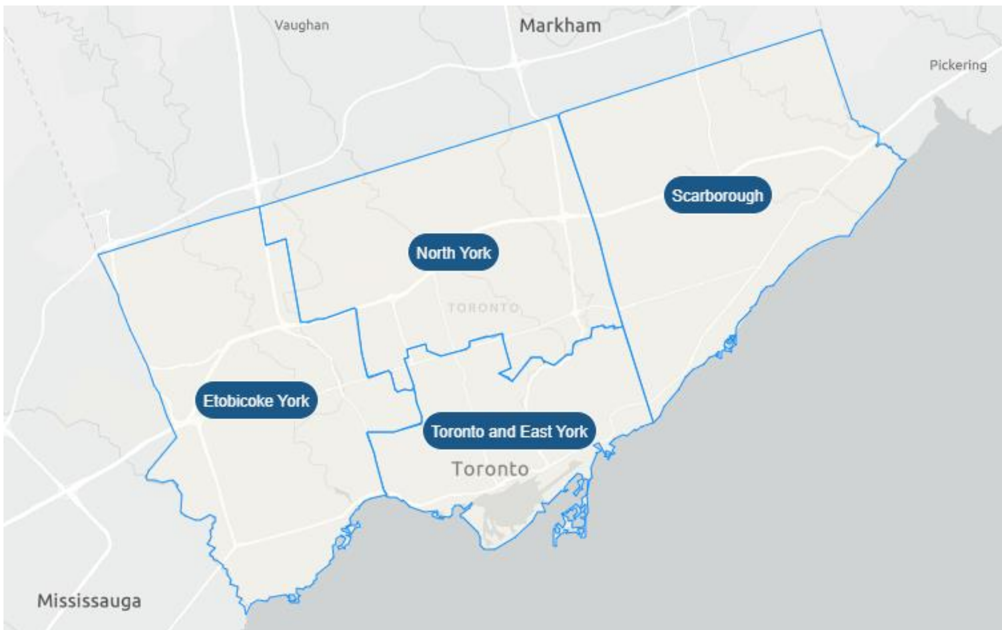
Figure 1: Visual of online survey scenarios with different attribute levels

3.2 Data collection

The next step after the development of the survey instrument is to determine a methodology to collect data. From prior studies, it is found that it is common for stated preference surveys to go through a pilot study process with a controlled group of participants to test the survey instrument (Perdomo, Rezaei, Patterson, Saunier, & Miranda-Moreno, 2014). As a result, the survey is first phased out in a pilot form with respondents that primarily include students in the University of Toronto. From the results and from general feedback from the pilot participants, minor changes are made to the survey instrument. In addition, the responses become a basis for any hypotheses that can be made for the final phase of data collection (See **Section 4.0**).

The survey instrument is, then, distributed to a representative sample after completion of the pilot phase with a help of a market research firm, Canadian Viewpoint. The geography of the respondents is restricted to the City of Toronto districts as the scenarios in the study is looking to gain insight on pedestrian experience within the city (See **Figure 2**). In total, 686 responses are collected before data cleaning. After data cleaning, 600 responses are kept for data analysis. The data cleaning process includes: eliminating data points in which respondents took less than approximately 300 seconds (or 5 minutes) to complete the survey (the actual minimum time on the database is 293 seconds), eliminating some blanks in the data due to respondents starting but not completing the survey, and eliminating any points in which the residents' place of residence was outside the City of Toronto boundaries. It is also important to note that the 600 responses have a demographic breakdown mirroring the City of Toronto 2016 census profile. This is done to ensure that the data sample of 600 is representative to the City of Toronto's entire 2016 population. **Appendix I** below shows a side-by-side comparison of the census profile and the survey respondents' profile.

Figure 2: City of Toronto boundary districts and boundaries for survey administration



3.3 Multinomial Logit Model and Mixed Logit Model

The statistical analysis of respondent's choices from the SP survey is explained using discrete choice modelling. A logit model is used to explain discrete choices among several alternatives with the assumption that they are all mutually exclusive. Within the model, random utility theory states that the decision maker n choses an alternative i given that it provides them with the highest utility (Wittink, 2011). While an individual's utility is specific to the individual, researchers can come to conclusions on the utility function through analysis of decision-makers' choices. The utility function is typically represented in the form shown in Eq (1).

$U_{ni} = V_{ni} + \varepsilon_{ni} \quad \forall i \quad (1)$

Here, the U_{ni} is the utility of alternative i within a choice set for individual n. V_{ni} is the systematic utility which can be explained with a set of explanatory variables, x_{ni}, and parameters, β = [$\beta_1 + \beta_2$, …. , β_K], as shown in Eq (2).

$V_{ni} = \beta x_{ni}\ \forall i = 1, \ldots, n$ (2)

ε_{in} is the error term that is unknown and explains any other factors or variables that can explain an individual's utility. The error term is assumed to be independently and identically distributed (IID) extreme value for the multinomial logit (MNL) model. A MNL is a simpler model as it yields closed-form equations. The equation that explains the probability that a person n chooses alternative i is shown in Eq (3) below (Train, 2002).

$$P_{ni} = \frac{e^{Vni}}{\sum_j e^{Vnj}} \quad (3)$$

While this form is simple, a major shortcoming of the MNL model arises from its assumption regarding the error term's IID property. This implies that the introduction or improvement of any alternative have the same proportional impact on the probability of each other alternative (Wittink, 2011). More specific to the case of this survey, the simple MNL model fails to look at observations involving more than one response from the same person, and fails to account for random tastes (differences in β) (Cascajo, Garcia-Martinez, & Monzon, 2017). A model that works to overcome these restrictions is the Mixed Multinomial Logit (MMNL) model. In contrast to MNL model where the coefficients β are fixed for everyone, the MMNL model allows for a vector of random coefficients β_n that vary across individuals (Perdomo, Rezaei, Patterson, Saunier, & Miranda-Moreno, 2014). The utility function for person n for alternative j in choice situation t can be presented in Eq (4).

$U_{njt} = V_{njt} + \varepsilon_{njt}$ with V_{njt} = $\beta_n x_{njt}$ (4)

Here, the error term, ε_{njt}, is IID extreme values over time, people and alternatives. MMNL probabilities are integrals of standard logit probabilities over a density of parameters. The integration of the product of logit formulas over all values of β makes it so that the correlation of errors across the choices of a given individual are captured (Perdomo, Rezaei, Patterson, Saunier, & Miranda-Moreno, 2014). The probability equation is expressed in the form shown in Eq (5).

$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta$ (5) where L_{ni} is the standard logit probability taken as a product of logit formulas evaluated at parameters β (see Eq (6)), and $f(\beta)$ is a density function (Train, 2002).

$$L_{ni}(\beta) = \Pi_{i=1}^{T} \left[\frac{e^{\beta_n x_{nit}}}{\sum_{J=1}^{J} e^{\beta_n x_{nJt}}} \right] (6)$$

The mixed logit model probability is a weighted average of the logit formula at different random values of β with the density, $f(\beta)$, defining the weight also called the "mixing distribution" (Train, 2002). As a result, there are two sets of parameters entering the mixed logit model: the parameter β that enter the logit model with a density $f(\beta)$, and the parameters that describe the density. If β is normally distributed with a mean b and covariance W, then b and W are the parameters describing $f(\beta)$ (Train, 2002). This study utilizes this concept by assuming a normally distributed β, and using the mean and variance to describe its density in the utility function. This concept is applied in this study with inclusion of only one random parameter with a mean and variance in each utility function. The choice of which parameter will be the "random" once is determined based on tests and outcome of statistical analysis. In addition, since respondents are asked to make repeated choices within a specific period, panel effects are added to the model on the same random parameter.

3.4 Data analysis

The data collected from the survey instrument is processed using an open-source choice modelling software called BIOGEME. To assess individual choice, the attribute levels are organized in a database in three sections (3 sections for 3 ranks). Every time an attribute level is chosen as a specific rank, it is assigned "1" vs "0" if it is not chosen. In order to compare attribute levels and their effects on one another, a dummy variable analysis is required. From hypotheses and pilot study results, the attribute level that is considered the least preferred is chosen as a base variable for comparison with others. **Table 2** below show the variables chosen as the base/reference for every attribute.

Table 2: Chosen reference variable for data analysis

	Through lane	Curb lane	Sidewalk
Chosen reference	car + car + car + car	none	1.6m pedestrian walkway + wall-side outdoor dining

Given 600 respondents and nine choice tasks (with 3 choices per task), there are a total of 5400 data entries for input into BIOGEME. For panel analysis, every 9 rows in the database corresponding to responses by one individual, is given one "ID" number that is ordered.

To assess different models and variable parameters as they are tested, there exists many statistical analysis tools. For this paper, the models are compared based on the following:

- Coefficient of parameters: the value/magnitude of the coefficient needs to be large enough that it shows that the variable is impacting the model. In addition, the sign of the coefficient needs to make some intuitive sense based on hypotheses stated.
- T-test: t-test is used to compare the sample mean of the data with the null hypothesis. The value of a t-stat will determine whether the null hypothesis should be rejected. Assuming a two-tailed t-test since the signs of coefficients are not definitively known, the absolute value of t-stat must be greater than or equal to the critical value of 1.96 to ensure a significance at 95% confidence interval to reject the null hypothesis.
- Log likelihood: a likelihood is a function that can be used to compare the plausibility of parameter. It is often estimated a log likelihood for computational purposes in which the values are always negative. Models with higher log-likelihood are considered better models.
- Akaike Information Criterion (AIC): AIC is used to estimate the likelihood of a model to estimate observed values. It is especially useful to perform model comparisons – a good model is the one with minimum AIC among all other models (Mohammed, Naugler, & Far, 2015).

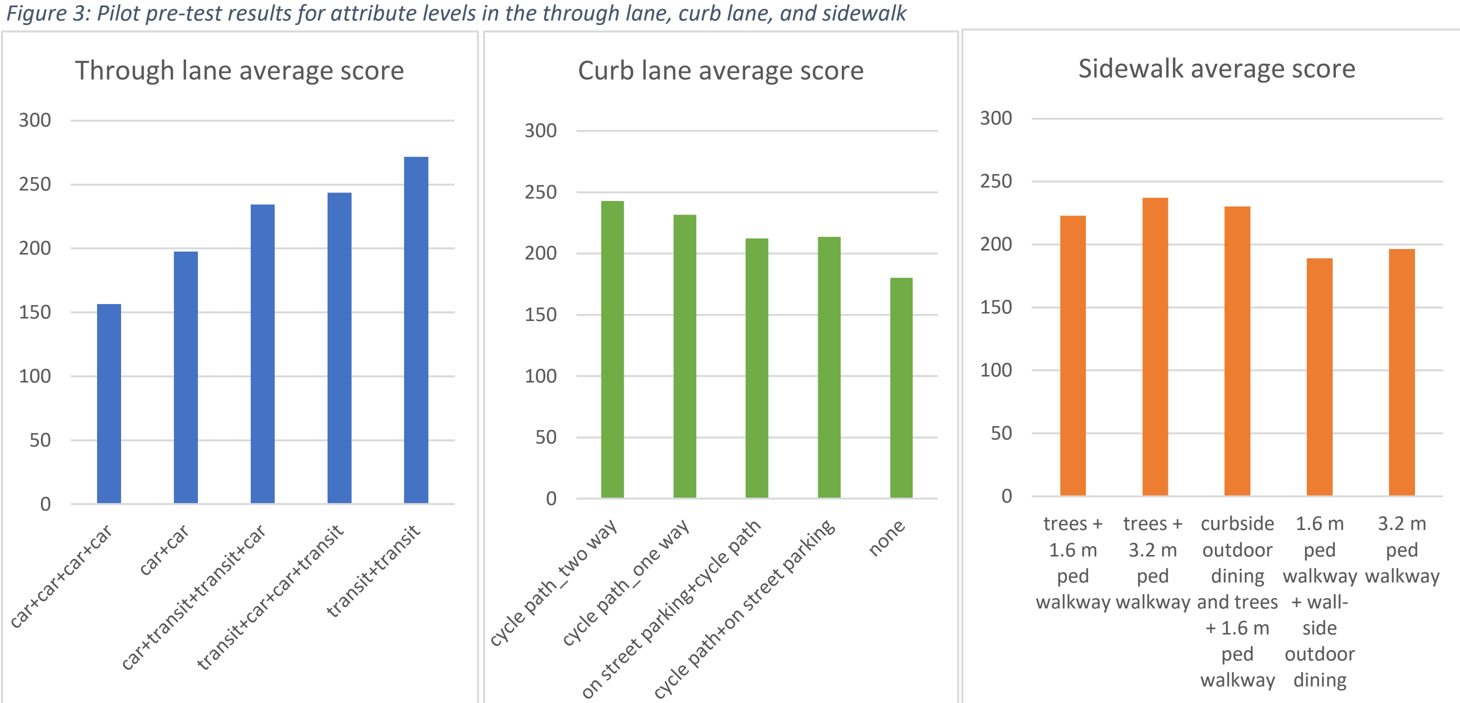
4.0 HYPOTHESES

The following section includes any precedence and any prior research to predict results for variables being assessed in the study.

4.1 Pilot study results

The pilot phase of the survey is used to determine if there are any patterns or predictions that can be made for the outcome of the real analysis. From the 109 responses collected, the number of times an attribute level is chosen as rank 1 is multiplied by 3, rank 2 by 2, and rank 1 kept as is (just the count) to retain average scores for every attribute level. The results are presented in **Figure 3**.

Figure 3: Pilot pre-test results for attribute levels in the through lane, curb lane, and sidewalk



For the through lane, the option with two transit-only lanes have the highest score while the option with two two-way car lanes (4 car lanes in total), is ranked lowest. The options with a mix of transit and car lanes also have high preference over the car lanes. For the curb lane, the highest score is attributed to the option with two-way cycle path with the “none” option given the lowest score. In the sidewalk, the option that includes trees and a 3.2m pedestrian walkway has the highest score with all options including trees ranked above options with no trees. These results and the general score can be predictors to test if the outcome is the same for the real study.

4.2 Other results

Very little information exists that measure perception of specific attribute levels in the through lane.

For the curb lane, there have been numerous surveys as well as location-specific guidelines regarding the use of curb lanes. Most polls pertain to perception of cycle path options. For example: a 2018 random survey of 800 Toronto residents reflected that more than 80% supported building protected bike lanes (Pelley, 2018). Similarly, the 2016 Bloor Street West Bike Lane Project included a feedback survey component post-installation which found that 74% of local residents backed the project. Among who drive and don't cycle, 57% opposed the lanes (Spurr, 2017). This, along with the results from the pre-test, support the hypothesis that the attribute levels with cycle paths may have greater utility over those without.

Additionally, a Copenhagen study looking at road safety and perceived risk of cycle facilities observed that streets with prohibited parking and cycle tracks brought on more accidents. It was found that streets with prohibited parking and cycle tracks prompted increase in turning traffic (in search for parking along side streets), and this consequently, led to an increase of accidents and injuries by 42% and 52%, respectively. A similar increase was also found for streets with cycle tracks and permitted parking but the increase in accidents and injuries were lower at 13% and 15%, respectively. In addition, the study found that in general, road sections with no parking observed a 24% increase in accidents whereas in sections with parking permitted, the accident rate fell by 14% (Jensen, Rosenkilde, & Jensen, 2006). While this study did not deal with perception, it reinforces the idea that an inclusion of parking on streets comes with some safety benefits that could form a positive perception for survey respondents.

Since the sidewalk element most pertains to pedestrians, prior studies on pedestrian perception can be found. A research paper from the journal of Korean Society of Civil Engineering (KSCE) did a comprehensive analysis of appropriate sidewalk width based on survey results, field sidewalk testing based on pedestrian path trajectory. The general perception showed that a wider sidewalk was always preferred (Kim, Choi, & Kim, 2010). However, this study did not include a combination of other elements paired with pedestrian walkway within the sidewalk that could influence choice (i.e. presence of outdoor dining patios). Preference of trees in sidewalks and their associated benefits have been well-documented. They are noted for having traffic-calming effect on local streets, improving air quality, and reducing urban heat island effect (City of Toronto, n.d.). These benefits are well-known and can be used to infer that scenarios with trees on the sidewalk will be ranked higher by respondents.

There exists a trade-off between narrower pedestrian walkway width and out-door dining options. Narrower sidewalk comes with issues of accessibility which can deter people from it. On the other hand, curb-side or wall-side outdoor dining options are often in locations that want to invite pedestrians to sit and enjoy the outdoors within the sidewalk. Little research has been done to compare the ramifications of this trade-off.

4.3 Expected results

Table 3 shows the expected results in terms of the sign of coefficients for all explanatory variables in comparison to the base variable from **Table 2**. The predictions are outcomes from the pilot study as well as external aforementioned research. These predictions are compared to the results section of the paper.

Table 3: Predictions of coefficient signs for explanatory variables

Attributes	name	Coefficient sign
transit + transit	TRAN	+
car + transit + transit + car	CT	+
transit + car + car + transit	TC	+
car + car	CARCAR	NS
two-way cycle path	TWOCYC	+
one-way cycle path	ONECYC	+
one-way cycle path + on street parking	CP	NS
on street parking + one-way cycle path	PC	NS
trees + 3.2m pedestrian walkway	TREE-WSW	+
trees + 1.6m pedestrian walkway	TREE-NSW	+
curbside outdoor dining and trees + 1.6m pedestrian walkway	SW-CURB	+
3.2m pedestrian walkway	SWONLY	+

NS = not significant

5.0 RESULTS

5.1 General Characteristics of Survey Respondents

The survey instrument includes questions regarding respondent demographic characteristics as well as the stated preference information. This section is a summary of these characteristics. Respondents' place of residence information as well as household size, household type, and dwelling type are shown in **Figure 4**, **Figure** , and **Figure 6**, respectively. The respondents' education and employment status, completed education, and income is shown in **Figure 5**, **Figure 6**, and **Figure 7**, respectively. Lastly, information on respondents' access to a driver's license, a personal vehicle, and cycling frequency is shown in **Figure 10**, **Figure 9**, and **Figure 8**, respectively.

Figure 4: Place of residence of respondents within City of Toronto districts

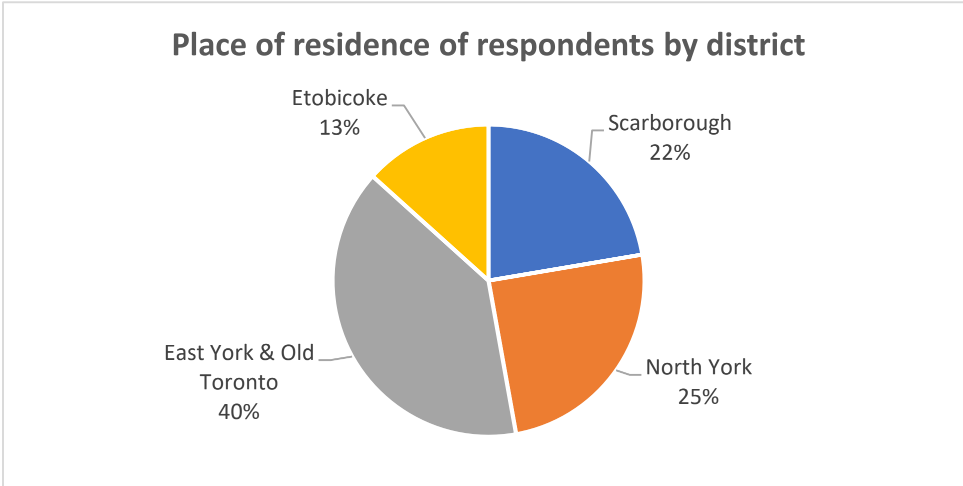
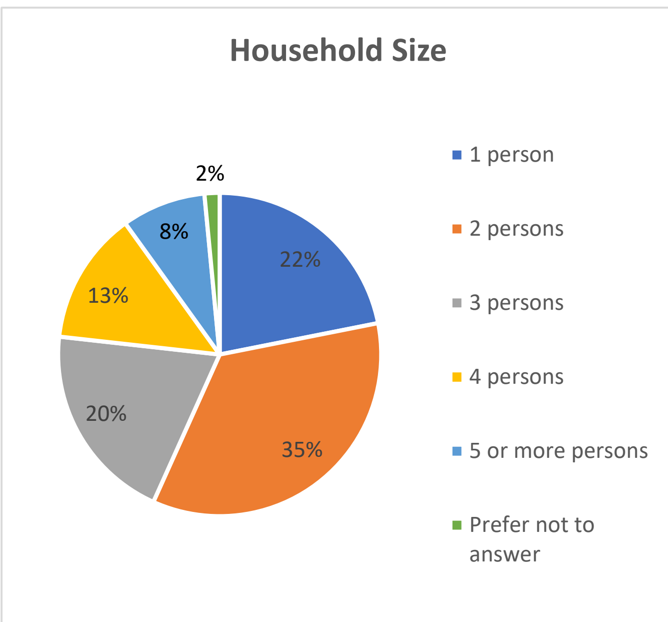


Figure 5: Survey respondents' household size and household type



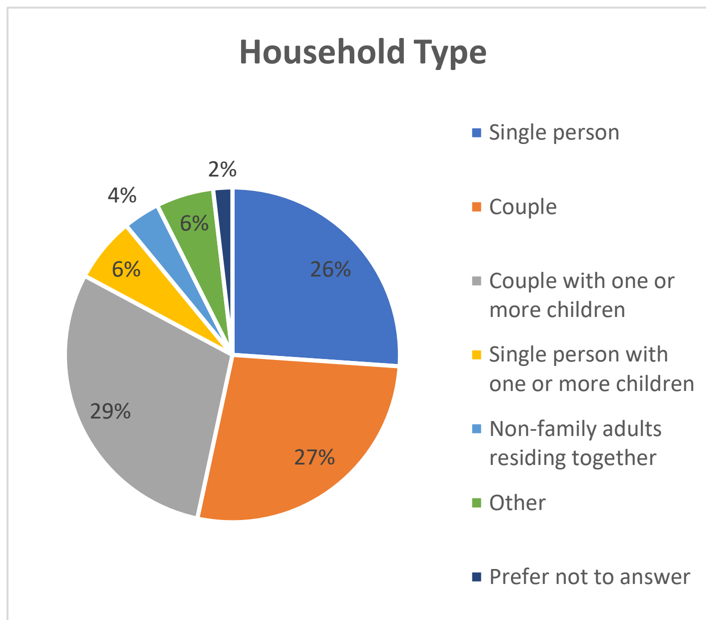


Figure 6: Dwelling type of survey respondents

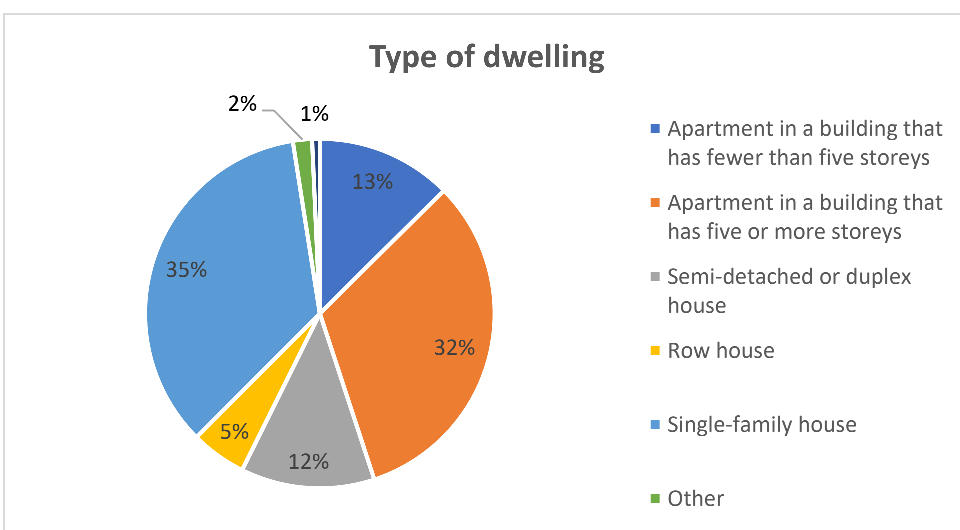


Figure 5: Education and employment status and completed of survey respondents

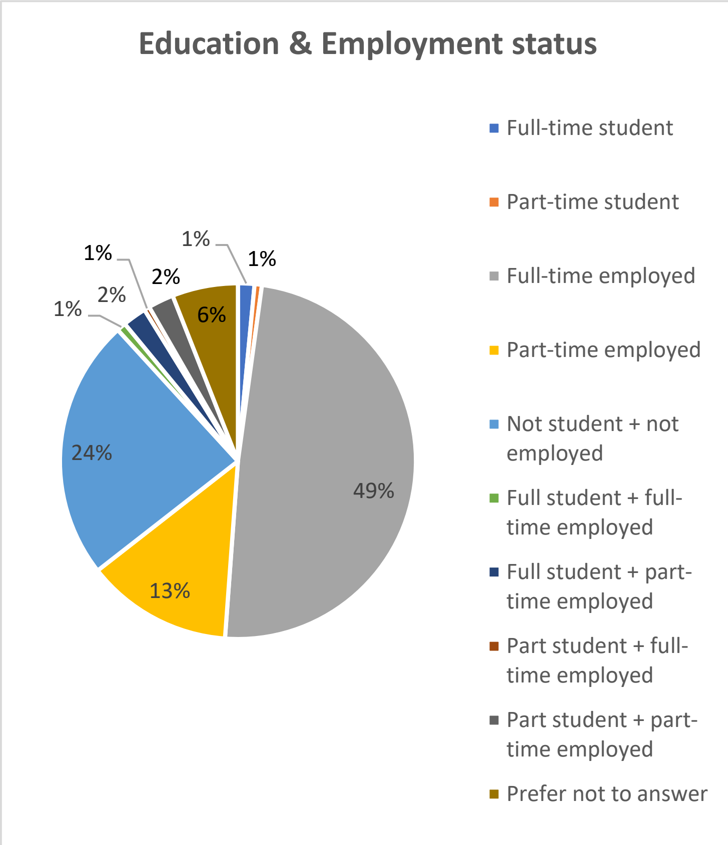


Figure 6: Education completed by survey respondents

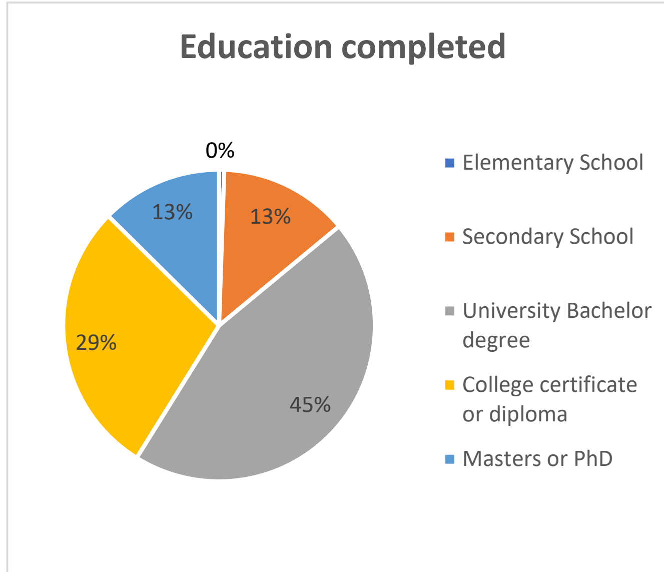


Figure 7: After tax income distribution of survey respondents

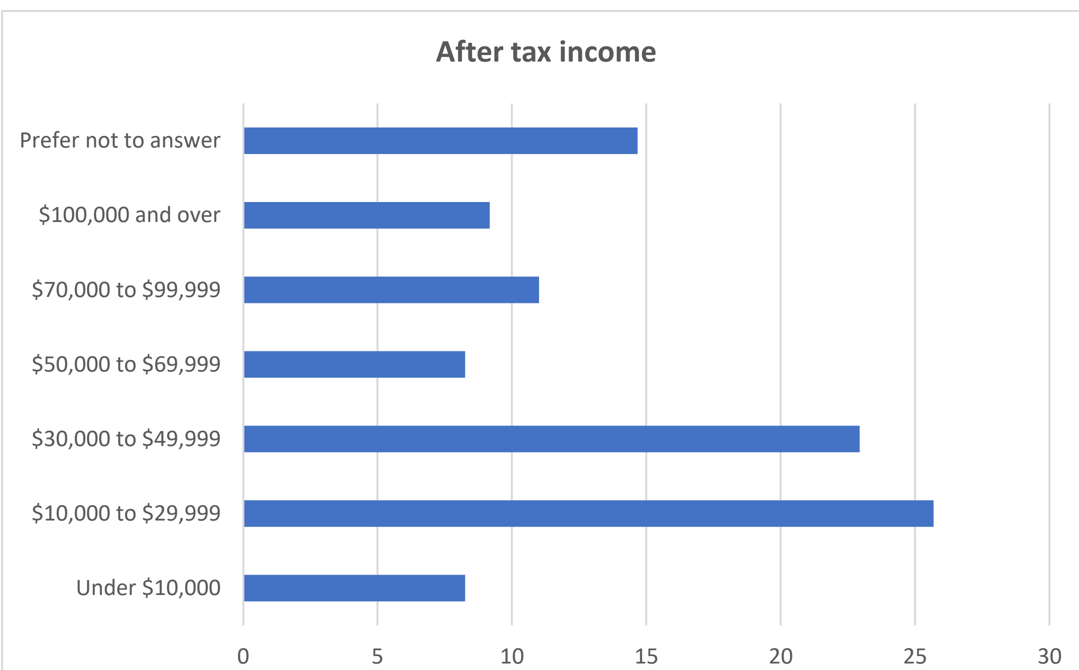


Figure 10: Survey respondents' access to a driver's license

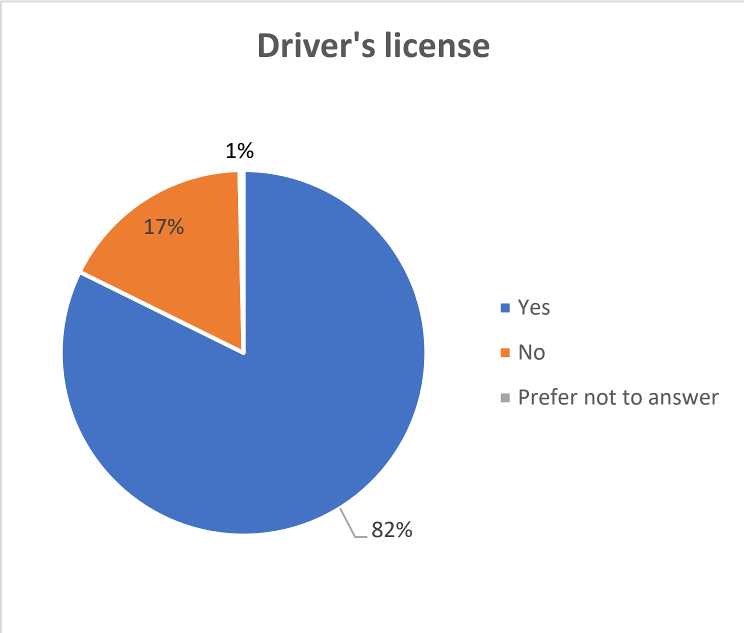


Figure 9: Survey respondents' auto-vehicle access

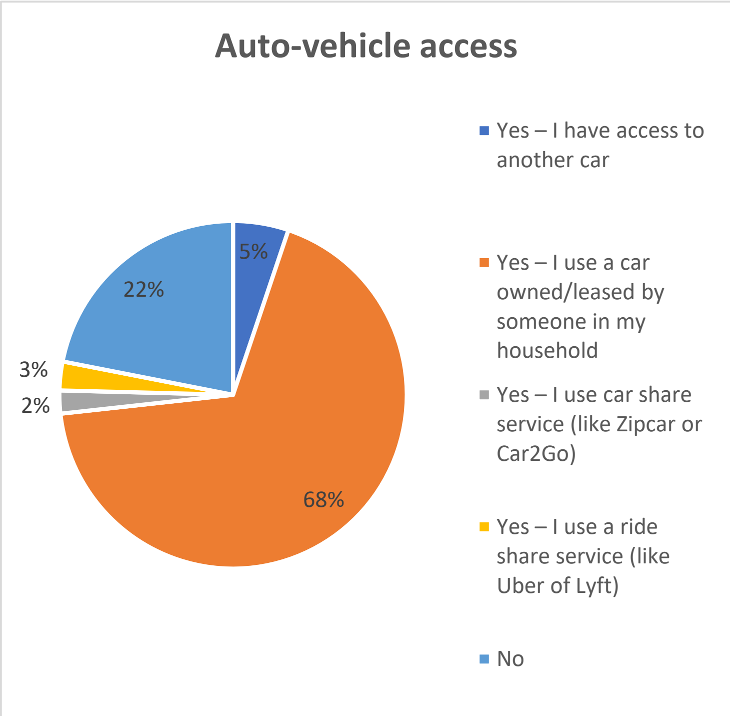
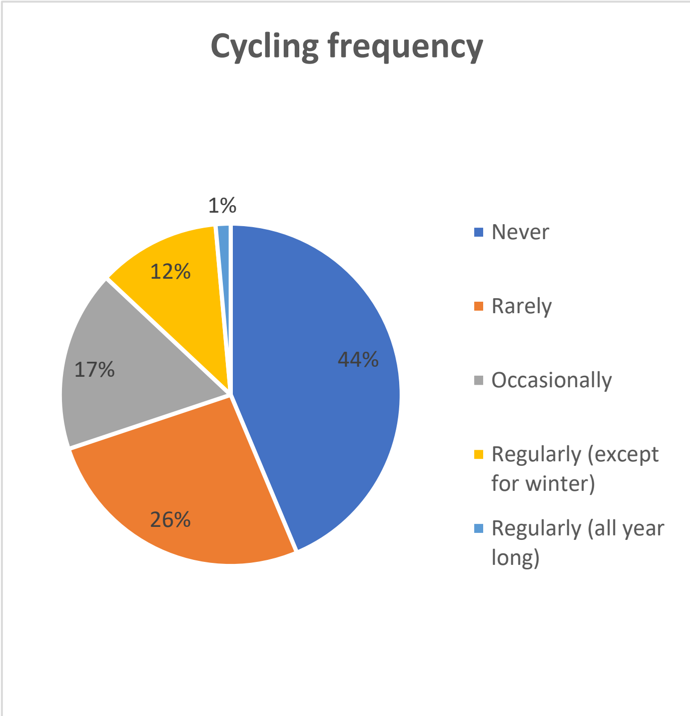


Figure 8: Survey respondent's cycling frequency



5.2 Model Results

Table 4 shows the final MMNL panel model estimated from the survey data. The mixed logit model is estimated with a random parameter for the variable "car + transit + transit + car" or CT showing its standard deviation coefficient. The value of this standard deviation parameter is high compared to its mean – indication that the coefficient does vary in the population (Train, 2002). The process of coming up with this model is done in three steps. Firstly, a simple logit model is estimated separately for every explanatory variable to assess its singular effect on the utility function (See

Appendix). Next, the variables that display a strong effect based on its coefficient, resulting t-test, log likelihood, and Akaike Information Criterion (AIC) is chosen and different combinations of these variables are tested to assess the combined impact on the utility function.

Appendix shows the process of this variable testing that lead to the final list of significant explanatory variables. Lastly, these chosen variables are tested using the simple MNL model. The MMNL model is chosen over the MNL model due to its more relaxed assumptions regarding the error term. In addition, **Table 5** shows that the MMNL panel model show better results (shown in green) with a higher log likelihood and a lower AIC value than the simple MNL model and a non-panel MMNL model with the same variable inputs.

Table 4: MMNL panel model results from survey data

Explanatory variables	Name	Coefficient (β)	t-test
car + transit + transit + car	CT	0.431	7.55
Standard deviation (car + transit + transit + car)	STDEV	0.808	11.86
transit + transit	TRAN	0.170	2.21
transit + car + car + transit	TC	0.451	9.84
two-way cycle path	TWOCYC	0.243	5.47
one-way cycle path	ONECYC	0.190	3.45
one-way cycle path + on street parking	CP	-0.325	-5.66
trees + 3.2m pedestrian walkway	TREE-WSW	0.305	7.11
curbside outdoor dining and trees + 1.6m pedestrian walkway	SW-CURB	0.185	4.20
3.2m pedestrian walkway	SWONLY	- 0.254	-4.97
Final log likelihood	-5702.58		
AIC	11423.16		

Table 5: MNL model results from survey data

Model type	Log likelihood	AIC
Multinomial logit	-5742.24	11502.48
Non-panel mixed multinomial logit	-5740.73	11499.5
Panel mixed multinomial logit	-5702.58	11423.16

6.0 Discussion

6.1 Significance of results

The following section explains the results from the previous section by comparing them to the hypotheses, as well as describing the quality of the model based on pre-determined statistical analysis tools.

6.1.1 Coefficients

The coefficients of the MMNL model confirm some of the hypotheses and results from the pilot study. All coefficients are significant (not a small value) and their sign reflects their preference when compared to the base variables defined in **Table 2**.

In the through lane, the coefficients for options with transit lane are positive, showing that alternatives with transit is preferred over alternatives with just car lanes. Additionally, the magnitudes of the coefficients show how much each of the transit alternatives are preferred over one another. From the data, it is seen that the option TC, inner transit lanes and outer curb-side car lanes, has the highest coefficient value with CT and TRAN following, showing its rank of preference. The results confirm the hypotheses of these variables being positively correlated with the utility function but provide more insight on how the preferences differ within alternatives. Lastly, the two car lanes option is missing as a variable in the model since its parameters were not significant when compared to the reference variable (four car lanes).

In the curb lane, two-way cycle path (TWOCYC) dominates as the preferred alternative with the highest coefficient value. The one-way cycle path (ONECYC) option is the next most preferred option. An interesting observation is that the variable CP (one-way cycle path inside, on street parking on curbside) has a negative sign for its coefficient. This shows that this alternative is preferred lower than an option with nothing on the curb lane (i.e. the reference variable).

In the sidewalk, the 3.2m walkway with trees (TREE-WSW) has the highest positive coefficient demonstrating its preference as an alternative over the rest of the alternatives and the reference variable (1.6m pedestrian walkway + wall-side outdoor dining). The curbside dining with 1.6m sidewalk (SW-CURB) also has a positive coefficient but its preference is lower than TREE-WSW. In contrast, the variable with just the 3.2m walkway (SWONLY) itself has a negative coefficient which denotes that its preference is lower than the reference variable (wall-side dining and 1.6m pedestrian walkway). This differs from the hypotheses and pilot results which gives a higher score for SWONLY.

6.1.2 T-test

All the t-test values pertaining to the model parameters are significant. Compared to the critical value of 1.96 for a two-tailed t-test, all parameters have absolute value of t-tests greater than this critical value. This means that the null hypothesis is rejected to show that the explanatory variables have significant influence over the utility.

6.1.3 Log likelihood & AIC

The results from **Table 4** shows the final model which has the highest final log likelihood value and the lowest AIC value when compared to earlier tests (**Appendix III**) and other tests carried out using simple multinomial logit modelling. This, further, attests to the quality of the final utility function and allows for a further level of confidence on the accuracy of the chosen parameters.

6.2 Limitations

This research paper outlines results with just the explanatory variables presented in the stated preference portion of the survey. To fully understand any patterns in the respondents' decision-making and choice, the socio-economic variables must be tested as well. This is a missing component in the model as none of the explanatory variables include these generic variables. Additionally, the data analysis does not include information attained from the revealed preference portion of the survey. This portion of the survey can provide a deeper understanding of what the respondents view to be a “good design” of streets based on what currently exists in Toronto, and as such, this missing information limits the full potential of the model. Further, the data analysis only considers scenarios that were ranked one and assigns that scenario to be the only available choice. However, the survey is designed for respondents to choose scenarios that are their first choice and their second choice. The point of time when they make this second choice, the first choice is no longer available and the number of scenarios is now limited to just two. This aspect is missing from this analysis.

The data from the survey is intended for the study of City of Toronto. The empirical results on pedestrian perception cannot be taken and applied for decision-making in other locations due to its context-sensitivity. The context-sensitivity nature of the survey is an advantage since the paper’s scope is to look at City of Toronto, but it limits its ability to be applied elsewhere since the decision-making and scenarios can be specific to the city.

Further limitations in the analysis arise from the mixed logit model itself. While it surpasses the assumptions restricting the simpler multinomial logit models, its assumptions regarding the normal distribution of the random parameter β may not be accurate. Moreover, the chosen model only considers effects of random parameter as applied to one variable, CT. While this is computationally easier, its application on other variables in combination might change results. In addition, the integration of the product of logit formulas to find probabilities does capture correlation effects of the error term but the error components will always come with uncertainty since it is aiming to explain factors affecting respondents’ choice outside of the measures that are outlined in the research.

6.2 Future Recommendations

The limitations of this study provide room for recommendations in the future. First, the data from the survey itself can be further utilized. This includes interacting socio-economic variables with the model to assess the impact on respondent characteristics’ and their effects on the preferences chosen. In addition, the revealed preference portion of the survey can be looked at to obtain a baseline or scale from which the SP data can be compared to – i.e. how the respondents feel about existing streets can provide some outlook on their choices on the SP portion of the survey. Further, adding another layer of data analysis considering the impacts of how the choices change when respondents are asked to choose between two scenarios when they make their second-ranking is recommended. These additions to the data analysis will present a clearer story regarding the preferences being studied.

This paper lays a foundation for studies on pedestrian preference within the City of Toronto. While its context-specificity is necessary to build the experimental SP design that is relevant to the respondents, the data collected is limited to Toronto. The methodology and approach of this study has the potential to be replicated for studying preference in other locations, and for studying pedestrian perception on different fields within transportation.

7.0 Conclusion

This research paper is addressing an observable lack of information and prior studies on pedestrian perception and experience within a streetscape. It provides a quantifiable measure on pedestrian preference of various street attributes through information gathered from an experimental SP survey. The design of the SP survey incorporates combining specific elements found within three attributes: through lane, curb lane, and sidewalk.

Survey data results are presented using mixed multinomial logit model with panel effects to account for multiple responses from each respondent at one time. The results show the following conclusions: there is a high preference of streets that include transit lanes as opposed to just car-exclusive lanes, with greater preference shown towards the alternative in which there is a mix of transit and car lanes as opposed to transit-only lanes; alternatives in the curb lane with exclusively just cycle paths are preferred while an option with a mix of inner cycle lane and curb-side parking lane has a lower preference over having no activity on the curb lane; interaction of sidewalk elements with tree components show a high preference among respondents while having a wider sidewalk without any activity has a lower preference than options with a narrower sidewalk with outdoor dining.

These results, with the survey sample representative to the City of Toronto population, provides insight on pedestrian preference for very specific design features found in streetscape. Noting that the current decision-making models when it comes to streetscape design has no component considering pedestrian perception as a variable, this study has taken steps to provide that layer of information. This is especially pertinent as much of the talks involving urban design guidelines increasingly feature active transportation components as well as more space allocated for pedestrian use and engagement. Understanding the pedestrian experience on streets can be a key component in determining how pedestrians, as stakeholders, contribute to the design themselves, and how the streets can be designed for maximum comfort.

8.0 References

Anciaes, P. R., & Jones, P. (2017). Estimating preferences for different types of pedestrain crossing facilities. *Transportation Research Part F (52)*, 222-237.

Audibert, M., He, Y., & Mathonnat, J. (2013). *Multinomial and Mixed Logit Modeling in the Presence of Heterogeneity: A Two-Period Comparison of Heathcare Provider Choice in Rural China.* Clermont-Ferrand: Centre d'Etudes et de Recherches sur le Développement International (CERDI).

Cascajo, R., Garcia-Martinez, A., & Monzon, A. (2017). Stated preference survey for estimating passenger transfer penalties: design and application to Madrid. *European Transport Reseach Review*.

City of Toronto. (2018, March). *Traffic Signal Vehicle and Pedestrian Volumes*. Retrieved from toronto.ca: https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/transportation/#7c8e7c62-7630-8b0f-43ed-a2dfe24aadc9

City of Toronto. (n.d.). *Street Design for Green Infrastructure*. Retrieved from toronto.ca: https://www.toronto.ca/wp-content/uploads/2017/11/90e0-Chapter-7.pdf

Gray, B. (2017, October 3). *Bloor Street West Bike Lane Pilot Project Evaluation.* Retrieved from https://www.toronto.ca/wp-content/uploads/2017/10/8ef6-cycling-bloor-backgroundfile-107582.pdf

Hensher, D. A. (1994). Stated preference analysis of travel choices: the state of practice. *Transportation 21* (pp. 107-133). Netherlands: Kluwer Academic Publishers.

Hui, N., Roorda, M., Saxe, S., Hess, P., & Miller, E. J. (2017). *Measuring the completeness of complete streets.* Transport Reviews.

Jensen, S. U., Rosenkilde, C., & Jensen, N. (2006). *Road safety and perceived risk of cycle facilities in Copenhagen.* Retrieved from https://www.vehicularcyclist.com/copenhagen1.pdf

Kim, S., Choi, J., & Kim, Y. (2010). Determining the Sidewalk Pavement Width by Using Pedestrian Discomfort Levels and Movement Characteristics. *KSCE Journal of Civil Engineering* .

Lusk, A. C., Filho, D. F., & Dobbert, L. (2018). Pedestrian and cyclist preferences for tree locations by sidewalks and cycle tracks and associated benefits: Worldwide implications from a study in Boston, MA. *Cities*.

Mohammed, E. A., Naugler, C., & Far, B. H. (2015). Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics. In *Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology* (pp. 577-602). Calgary.

Pelley, L. (2018, July 31). *80% of Toronto residents support building protected bike lanes, poll finds.* Retrieved from CBC: https://www.cbc.ca/news/canada/toronto/bike-lane-poll-toronto-1.4766745

Perdomo, M., Rezaei, A., Patterson, Z., Saunier, N., & Miranda-Moreno, L. F. (2014). Pedestrian preference with respect to roundabouts - A video-based state preference survey. *Accident Analysis and Prevention 70*, 84-91.

Project 2.1 Evaluation of Complete Streets. (2017). Retrieved from University of Toronto Transportation Research Institute: https://uttri.utoronto.ca/research/projects/icity/papers/theme-two/project-2-1-evaluation-complete-streets/

Spurr, B. (2017, October 11). *Bloor bike lanes should stay: city report*. Retrieved from The Star: https://www.thestar.com/news/gta/2017/10/11/bloor-st-bike-lanes-should-stay-city-report-finds.html

The City of Calgary. (n.d.). *Cycle Track Design Best Practices.* Retrieved from https://www.placespeak.com/uploads/assets/Cycle_track_network_study-_design_1.pdf

Train, K. (2002). *Discrete Choice Methods with Simulation.* Cambridge University Press.

Wittink, L. T. (2011). *Choice modelling: an overview of theory and development in individual choice behaviour modelling.* BMI Paper.

Appendix I

2016 Census profile vs. survey respondent profile

Table 6: Census profile vs survey profile based on age and gender

Age group	2016 census profile		Survey profile	
	% male	% female	% male	% female
20-29 years	49	51	47	53
30-39 years	48	52	44	56
40-49 years	48	52	45	55
50-64 years	48	52	48	52
65 years and over	43	57	42	58

Table 7: Census profile vs survey profile based on geography

Place of residence	2016 census profile (%)	Survey profile (%)
East York & Old Toronto	37	40
Etobicoke	15	13
North York	25	25
Scarborough	23	22

Appendix II

Univariate variable testing

Variables	Test	1	2	3	4	5	6	7	8	9	10	11	12
car+car	β	-0.212											
	t-test	-5.2											
car+transit+ transit+car	β		0.215										
	t-test		5.76										
transit+car+ car+transit	β			0.243									
	t-test			6.29									
transit+transit	β				0.49								
	t-test				7.85								
two-way cycle path	β					0.25							
	t-test					6.79							
one-way cycle path	β						0.0654						
	t-test						1.56						
on street parking + one-way cycle path	β							-0.17					
	t-test							-4.03					
one-way cycle path + on street parking	β								-0.44				
	t-test								-8.49				
trees + 1.6m pedestrian walkway	β									-0.027			
	t-test									-0.73			
trees + 3.2m pedestrian walkway	β										0.237		
	t-test										6.71		
curbside outdoor dining and trees + 1.6m pedestrian walkway	β											0.17	
	t-test											4.36	
3.2m pedestrian walkway	β												-0.32
	t-test												-7.41

Appendix III

Combined variable testing

Variables	Test	1	2	3	4	5	6	7	8
car+car	**β**		0.034						
	t-test		0.70						
car+transit+transit+car	**β**					0.288	0.482	0.493	0.460
	t-test					7.37	10.82	11.01	10.46
transit+car+car+transit	**β**						0.442	0.467	0.435
	t-test						9.64	10.05	9.67
transit+transit	**β**	0.463	0.488	0.365	0.383	0.305	0.167	0.161	0.184
	t-test	7.42	6.77	5.72	5.89	4.43	2.18	2.10	2.42
two-way cycle path	**β**				0.232	0.229	0.301	0.317	0.244
	t-test				6.29	6.11	6.45	6.8	5.58
one-way cycle path	**β**						0.263	0.265	0.182
	t-test						4.66	4.67	3.36
on street parking + one-way cycle path	**β**						-0.0113	0.0421	
	t-test						-0.22	0.80	
one-way cycle path + on street parking	**β**	-0.426	-0.435	-0.46					-0.313
	t-test	-8.11	-8.04	-8.68					-5.61
trees + 1.6m pedestrian walkway	**β**								
	t-test								
trees + 3.2m pedestrian walkway	**β**			0.217	0.184	0.169	0.240	0.299	0.297
	t-test			5.85	5.01	4.47	6.02	7.06	7.07
curbside outdoor dining and trees + 1.6m pedestrian walkway	**β**							0.199	0.181
	t-test							4.47	4.17
3.2m pedestrian walkway	**β**					-0.307	-0.303	-0.227	-0.232
	t-test					-6.48	4.66	-4.36	-4.66
Final log likelihood		-5866.8	-5866.6	-5849.8	-5869.6	-5828.5	-5767.9	-5757.94	-5742.2
AIC		11733.6	11733.1	11699.7	11739.2	11657.01	11551.75	11533.87	11502.48