UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING
Transportation Research Institute

UTTRI

WEB TOOL DESIGN FOR HOUSEHOLD TRAVEL SURVEYS

2016-11-04

Transportation Tomorrow Survey 2.0

Brittany Chung, Siva Srikukenthiran

Khandker Nurul Habib & Eric J. Miller

EXECUTIVE SUMMARY

The Transportation Tomorrow Survey (TTS) is a household travel survey of the Greater Golden Horseshoe Area (GGHA) that has been undertaken every five years since 1986. In the past, the TTS survey was conducted through Computer-Assisted Telephone Interviews (CATI). Due to recent advancements in internet and smart phone technologies, land-line phone use has been declining, which has presented challenges to the methods of data collection used by the TTS. A review of the literature and the experience of UTTRI researchers revealed that web surveys have the potential to replace and/or complement the CATI method for household travel surveys. A TTS 2.0 review paper conducted in 2015, "Current State of Web-Based Survey Methods", explored and discussed the potential uses of web surveys in the new design of the TTS. The report covered sampling frame issues faced by web surveys, compared web and traditional survey modes, investigated the potential of multi-day web surveys, and analyzed differences between web and CATI respondents. This report acts as a follows-up, as well as an extension to that report.

The key objective of this report was to investigate design considerations in the literature that could be implemented in a web version of the TTS. The purpose of the research was to identify methods to achieve improved accuracy and completeness of self-reported information using a web-based interface, and to identify critical gaps in a literature that could be evaluated and tested in the first round of field tests for the TTS 2.0 project to occur in summer of 2017. Four specific areas were investigated: general user design considerations, design of web survey questions, how to design for accurate and complete responses, and design of interactive web-based maps.

The user interface of a website is the primary line of communication between a surveyor and the respondents, and if not designed well, can result in frustration, aggravation and stress for the survey taker. Various studies in the literature have proven that the user interface of a web survey can significantly influent respondents' attitudes of the survey, having a direct impact on data quality and response burden. To inform a new web-version of the TTS, literature on several areas of web design (screen navigation, page layout, colour, media and typography) were investigated looking for best practices, and guidelines are presented for each area. Furthermore, web accessibility guidelines which are legally required by the Ontario government were outlined in the report to ensure the compliance of the new TTS web survey.

A key process in the design of web surveys is planning the layout and presentation of the questions. The layout and presentation of a survey can influence the respondent's interpretation of the questions. Thus, it is important to design for clear, comprehensive survey questions to avoid misinterpretations and unnecessary burden on the respondents. As with the user interface design, best practices for question design were developed based on an examination of the literature. Elements presented include how to word and structure survey questions, design and organization and use cases of a variety of multi-select structures (radio buttons, multi-select checkboxes, dropdown selects, matrices), text boxes, and ranged-input (sliders, Likert scales).

While following proper design and proper use of web elements can help achieve accurate and complete responses, elements outside of the questions themselves can play a key role in data accuracy. Several of these are presented, including how best to introduce the survey to potential respondents, best ways to indicate their progress to respondents while taking the survey, the need for providing definitions and clarifications to avoid ambiguity, and allowing for optional responses where

possible. Additionally, the issue of proxy bias, where one member of a household reports on trips for other members, was also closely examined, given that it can compromise quality and completeness of responses. The effect of proxy bias on past travel studies, the demographics of proxy respondents, and methods to correct and avoid bias (mainly with trip rate adjustment factors) are presented. Data quality can also differ between prompted recall and announce-in-advance surveys. Prompted recall surveys ask respondents to recall what happened on a prior day, while the in the latter method, the surveyor announces to the respondents ahead of time that they must report their trips for a specified date in the future. Although this is not a very active area of research, a literature review of the two methods are presented in this report.

Finally, the use of interactive web-based maps is examined, given the importance of collection of trip data and the high burden in collecting this type of detailed data using traditional CATI methods. Prior studies detailed are the Utah Travel Survey, the Toronto-based StudentMoveTO, the U.S. National Household Travel Survey, and the Edmonton and Region Household Travel Survey; specifically, inspirations that can be taken from these prior efforts are discussed. The two main map interfaces available, single point location and trip routing, are discussed, including analysis of how users in prior surveys have utilized these input methods. Guidelines are provided on how best to design map-based input for travel surveys.

With input from this extensive literature review of website, survey and map-input design, an in-house web-based survey platform was developed, and a version of the TTS was reproduced. In order to both evaluate the developed tool, and guide improvements, two main field tests are recommended for the summer of 2017. The first will be a series of small focus groups, where the focus will be user experience testing to gain a detailed understanding of how respondents use and navigate the survey, more specifically the map interface trip questions. The second will be a large-scale test on prior TTS respondents to examine the profile of respondents, including demographics and completion times and rates. Also to be studied are two methods of survey administration, comparing mixed self and proxy reporting (where multiple household members can be interviewed without increasing burden), and analyzing data quality differences between prompted recall and announce-in-advance surveys. Results will be presented in the fall of 2017.

WEB TOOL DESIGN FOR HOUSEHOLD TRAVEL SURVEYS

TRANSPORTATION TOMORROW SURVEY 2.0

Table of Contents

List of Tables

List of Figures

1 INTRODUCTION

The Transportation Tomorrow Survey (TTS) is a household travel survey of the Greater Golden Horseshoe Area (GGHA) that has been undertaken every five years since 1986 (DMG, 2016). The survey is designed and executed through the collaborative efforts of local and provincial agencies to collect travel data that contribute to transportation planning and investment decisions in the region. In the past, the TTS survey was conducted through Computer-Assisted Telephone Interviews (CATI). Due to recent advancements in internet and smart phone technologies, land-line phone use has been declining, which has presented challenges to the methods of data collection used by the TTS. In response, the TTS 2.0 project was created to develop the next generation of travel survey methods for the region, with a goal of overcoming the current and foreseeable challenges the TTS has or may face.

A review of the literature and the experience of UTTRI researchers revealed that web surveys have the potential to replace and/or complement the CATI method for household travel surveys. A TTS 2.0 review paper conducted in 2015, "Current State of Web-Based Survey Methods", explored and discussed the potential uses of web surveys in the new design of the TTS. The report covered sampling frame issues faced by web surveys, compared web and traditional survey modes, investigated the potential of multi-day web surveys, and analyzed differences between web and CATI respondents (Loa, et al., 2015). This report acts as a follows-up, as well as an extension to that report.

The key objective of this report was to investigate design considerations in the literature that could be implemented in a web version of the TTS. The purpose of the research was to identify methods to achieve improved accuracy and completeness of self-reported information using a web-based interface, and to identify critical gaps in a literature that could be evaluated and tested in the first round of field tests for the TTS 2.0 project. Section 2 identifies web interface design elements proven in the literature that could improve the usability of the TTS web survey. Section 3 discusses the design of various web survey questions (i.e. radio buttons, text boxes, sliders, etc.) that could be used to collect household data. Section 4 identifies key web survey elements and methods that directly and/or indirectly influence the accuracy/completeness of survey responses; it also details how they could be properly applied to the future TTS web survey. Section 5 explores the use of interactive maps in household travel surveys, and design considerations for the collection of trip data. Finally, Section 6 proposes field tests based on the gaps identified in the research presented in this report.

2 USER INTERFACE DESIGN CONSIDERATIONS

The user interface of any web survey is the medium upon which the user and the website/application program interact, and is the primary line of communication between the surveyor and the respondents. Much like any line of communication, such as a conversation through a telephone or an e-mail, any confusion or disconnect can cause frustration, aggravation, and stress. This logic also applies to poorly designed user interfaces. Various studies in the literature have proven that the user interface of a web survey can significantly influence respondents' attitudes toward the survey. The design can be the difference between the respondent's acceptance or rejection of the survey. Poor web survey designs are confusing, inefficient and can evoke negative emotions for users. Negative attitudes towards a survey may result in higher chances of the respondent making mistakes and cause respondents to lose motivation to complete the survey with care and honesty. Ultimately, a poor design can compromise the response rate and quality of the data collected.

It is apparent that the user interface of a web survey plays a critical role in the survey's success and/or failure. The interface of a future web-based version of the TTS should be carefully designed for success; this involves designing for ease of use, clear navigability, and low respondent burden (DMG, 2015). This section of the report attempts to identify web interface design elements proven in the literature that can improve the usability of a web survey. The purpose of the literature review is to understand general concepts about human-computer interaction (HCI) and how a user-friendly design can be achieved in the future TTS web survey.

2.1 Screen Navigation and Flow

New user interface designers often fail to realize that very few users, if any at all, diligently read every single line of text on their web page, regardless of how well-crafted this text is. Instead, users typically scan pages until something peaks their interest, or until they have found that for which they had been searching (Galitz, 2002; Krug, 2006). Understanding the scanning behavior used to navigate through websites/web surveys is critical in the creation of a user-friendly design. A successful web survey interface should conform to the natural behaviours of web-users to ease the user's experience.

For Western readers, the eye is initially anchored to the top left corner of a paper, as this is where a heading or a paragraph begins. Various user interface designers have argued that web users also adopt this anchor point when scanning web pages and browsing. An eye-tracking study conducted by Nielson (2006) also supported this statement. Nielson (2006) tracked the eye movement patterns of over 200 web users viewing a various range of web pages. The study strongly reinforced the idea that exhaustive reading is rare for web-users. His study also concluded that the dominant scanning pattern is from left to right and then top to bottom of the page (Nielson, 2006). The scanning of the page horizontally to the right becomes progressively shorter as the user moves down the page (Nielson, 2006). Heat maps of Nielsen's eye-tracking study illustrating the scanning pattern described are provided as Figure 1. As shown in the heat maps, the scanning pattern appears in the shape of an "F"; thus, Nielson (2006) coined this scanning behaviour as the "F-Pattern".

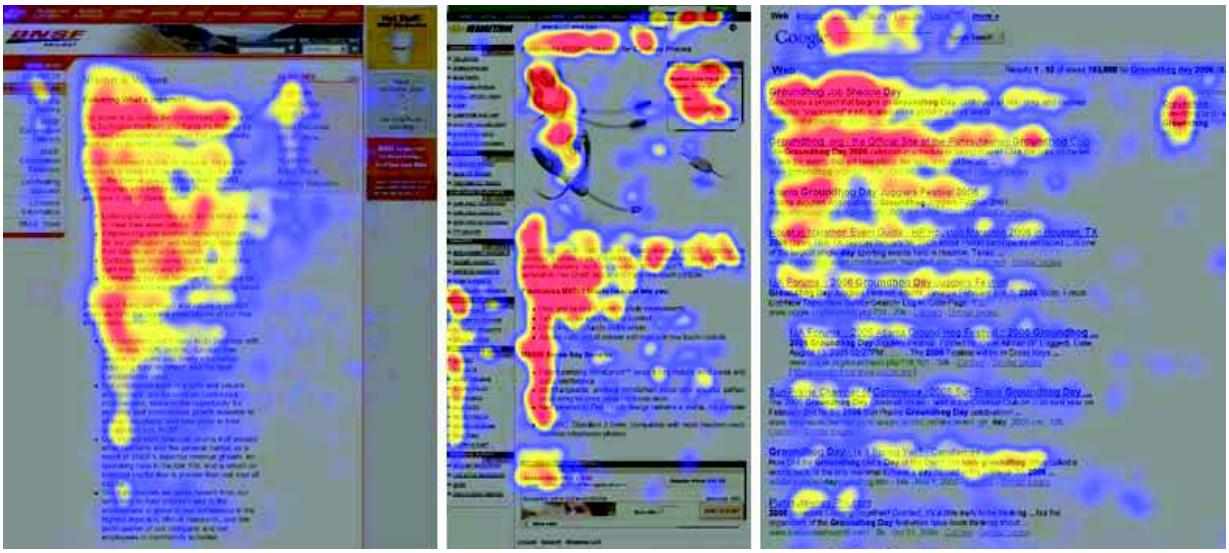

FIGURE 1 – HEAT MAPS FROM NIELSEN (2006) EYE-TRACKING MOVEMENT STUDY ILLUSTRATING THE "F-PATTERN"

It is evident from the users' scanning pattern that the top left corner of a web page is most visible to the user while the bottom right is least visible. Therefore, Nielsen (2006) recommends that the first paragraph should include the most important information. A study by Wilkinson and Payne (2006) also arrived at the same recommendation. Their study revealed that web-users employed a 'skimming by satisficing' strategy rather than reading the entire text (Wilkinson & Payne, 2006). The participants tended to look at the beginning of a paragraph/text section, followed by skimming the remainder of the text to decide whether to read or skip to the next section (Wilkinson & Payne, 2006). Based on this pattern of scanning through a web page, Wilkinson and Payne (2006) suggested that web-users typically do not read more than the first paragraph or text section of a web page.

Furthermore, Nielsen (2006) suggested that the designer should make use of headers, subheadings, and bullet points to draw user's attention to information placed further down the page and to guide the F-pattern scanning behaviour as the user scrolls down the page. The direction of movement along the web page should be obvious, consistent, and have a rhythmic flow that follows the 'F-pattern'. The headings should also be as descriptive as possible in the beginning because the users are more likely to read the first few words of the heading than the words near the end of the line.

Shortly following Nielsen's discovery of the "F-pattern" scanning behaviour, Shrestha and Lenz (2007) conducted a similar study on Wichita State University students, with ages ranging between18 to 26 years old. The conclusions of their study supported Nielsen's "F-pattern" theory, namely that users' eyes tend to fixate at the beginning of a line more than the end of a line, and that the users fixated on the first few lines longer than for subsequent lines down the page (Shrestha & Lenz, 2007). However, their study found that the "F-pattern" scanning pattern did not hold true when a user was browsing or searching through a picture-based webpage (Shrestha & Lenz, 2007). Use of images and pictures in web design will be further discussed in Section 2.5 of this report.

Although Nielsen, Wilkinson and Payne, and Shrestha and Lenz argued that the most visible section of a web page is the near the top, there was a popular phenomenon called "banner blindness" that

suggested there was an exception to this theory. It is believed that web-users quickly get accustomed to the header at the top of a web page, and will ignore any information that it contains. A later study conducted by Nielsen (2007) recorded the eye movement of 232 users on a variety of websites; it was found that users almost never looked at banner headings or anything that resembled an advertisement, even though it might be at the very top of a web page. Figure 2 presents heat map drawings from Nielsen (2007) eye-tracking study. The heat maps display three viewing scenarios: quick scanning (left image), partial reading (centre image), and thorough reading (right image).

FIGURE 2 - HEAT MAPS FROM NIELSEN (2007) EYE-TRACKING MOVEMENT STUDY ILLUSTRATING THE "BANNER BLINDNESS"

A more recent study by Guner and Inal (2015) arrived at the same conclusion as Nielsen. The study examined university students eye movement patterns on a Turkish government website. The results of the study revealed that banners, including a picture of the head of the public institution, was often ignored by the students (Guner & Inal, 2015). Nearly half of the students recalled where the banner was placed, but none of the students could recall the information on the banner (i.e. information about the institution head, picture, social media information…etc.) (Guner & Inal, 2015). This banner blindness phenomenon suggests that banner headings should be used carefully such that they do not resemble advertisements, and that important and detailed information should not be presented in the banner.

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary of the recommended design considerations is presented in Table 1.

TABLE 1 – USER INTERFACE DESIGN CONSIDERATIONS RELATED TO SCREEN NAVIGATION AND FLOW

Design Considerations	Reasoning
Key information and most used action buttons should be placed near top left corner of the *web page*	• The top left corner of a web page is most visible to users and the bottom right is the least visible.
The first paragraph should include the most important information	• Exhaustive reading of a web page is rare of web-users. • Web users tended to look at the beginning of a paragraph/text section and then they would skim the remainder of the text to decide whether to read or skip to the next section. • Text read/scanned by the web-user is less than or equal to a paragraph or text section as suggested by the "F-pattern" theory.
Use headers, sub-headings and bullet points to draw user's attention to information further down the *web page*. Ensure to keep an obvious, consistent, and rhythmic format.	• Helps to guide the "F-pattern" scanning behaviour as the user scrolls down the page.
Headings should be as descriptive as possible in the beginning.	• Users are more likely to read first few words of the heading than the words near the end of the line.
Ensure banner headings do not resemble advertisements or contain key information.	• Users often overlook banner headings; this is referred to as the "Banner Blindness" phenomenon.

2.2 Scrolling vs Paging

A web survey can be presented in two forms: a scrolling design and a paging design. As the names suggest, a scrolling design presents the entire survey on a single web page while paging design sections the questionnaire into several web pages. Several studies that have compared the two designs have shown no significant difference in participation rates and/or drop-out rates. Peytchev et al.'s (2006) study conducted a web survey on drug and alcohol use for over 21,000 undergraduate students and it featured a scrolling design and a paging design. When comparing the survey results, they found no difference in response rates, drop-out rates, non-substantive answers (i.e. explicit refusal or "don't know" responses), nor significant distribution in the responses between the two designs (Peytchev, et al., 2006). However, the results did reveal that the scrolling design had longer completion times and missing responses (Peytchev, et al., 2006). Manfreda et al. (2002) conducted a similar experiment comparing the two designs and reported similar findings. Manfreda et al. (2002) argued that the scrolling design of a web survey increases the likelihood of the respondents unintentionally skipping questions and thus results in higher item nonresponse rates.

Both Peytchev et al. and Manfreda et al.'s web surveys were conducted over personal computers (PC). Web surveys on smaller screens, such as smartphones or tablets, may reveal different results regarding paging and scrolling designs. Mavletova and Couper (2014) completed a follow-up study

among panelists who were non-respondents or drop-outs of a large mobile web survey in Russia. The overall participation rate of the mobile study was 53.6% (Mavletova & Couper, 2014). The follow-up study involved 660 panelists, of which 156 were non-respondents and 504 were drop-outs of the large mobile web survey (Mavletova & Couper, 2014). The results revealed that the main reason for dropping out of the mobile survey was due to technical difficulties; 47% of the drop-outs argued that some response options were not immediately clickable (Mavletova & Couper, 2014). Others argued that the mobile survey had long download times (24%) and that they were experiencing slow Internet speeds (16%). McGeeny and Marlar (2013) also conducted an experiment comparing paging and scrolling web survey designs involving both PC and mobile web respondents but did not reveal any key differences between the PC and mobile results. However, they found that the scrolling design of a 13-question survey had higher participation rates and lower drop-out rates than the paging design (McGeeney & Marlar, 2013). It is likely that the illusion of a one-page scrolling design appears shorter to a respondent than a 13-page survey, although they present the same content.

The implications of scrolling and paging designs vary in the literature; however, there appears to be more evidence on the negatives of scrolling designs for lengthy questionnaires than for paging designs. As mentioned in the papers discussed, the scrolling designs had longer completion times and increased the likelihood of missing responses. Furthermore, paging designs facilitates the development of "templates" for generic types of questions. These "templates" are more easily integrated into a page of a custom-tailored survey design than into a scrolling design. Since efficiency and flexibility are key goals for questionnaire design, it is recommended that the future TTS web survey should adopt a paging design. Unfortunately, none of the papers provided guidelines on how many survey questions should be presented on a page. The survey designer should be aware that presenting fewer questions on a page will result in greater number of pages. As mentioned in McGeeny and Marlar's (2013) study, the number of survey pages can cause respondents to perceive the survey to be longer than its actual length. Since the a TTS-like web survey will be rather lengthy, it should find a balance in the structure of its paging design such that its perceived length does not discourage respondents.

2.3 Paged Web Survey Navigation

Web surveys should be forgiving and flexible to permit users to review, change, and/or undo actions when necessary. Paged web surveys commonly have a "Next", "Continue" or a forward arrow button that allows the respondents to proceed to the next page after they have completed the questions on the current page. It is believed that the "Next" button also gives the opportunity for respondents to check and reconsider their answers before proceeding (Tourangeau, et al., 2013). It has been found to be the most frequently used function in paged surveys and, thus, it should be the most visible button on the page (Faulkner, 1998; Leavitt & Sneiderman, 2004; Dillman, et al., 2009). The counter function, the "Previous" button, has been shown to be rarely used. Often the "Previous" button is used when the respondent has difficulty understanding the early question that routed them to a question that does not apply to them (Tourangeau, et al., 2013).

Several studies in the literature experimented with the placement of the "Next" and "Previous" buttons in web surveys to determine the ideal positioning of these key functions. Couper, Baker, and Mechling's (2011) study found that the position of the "Next" and "Previous" buttons had negligible effects on the breakoff rates or completion times of their survey. However, their results revealed that

respondents were more likely to unintentionally use the "Previous" button when it was to the right of the "Next" button (Couper, et al., 2011). They also found that when the "Previous" button was more prominently displayed, respondents often accidently clicked on the button, which lengthened their completion times (Couper, et al., 2011). Couper, Baker, and Mechling (2011) recommended reducing the prominence of the "Previous" button by either using a hyperlink instead or placing it below the "Next" button. The findings of an earlier study by Wroblewki (2008) were consistent with the results of Couper, Baker, and Mechling. Wroblewki (2008) additionally found that more frequently used action buttons should be placed on the left side of the screen. This was consistent with the F-pattern scanning pattern discussed in Section 2.1. A fairly recent study for the U.S. Census Bureau also experimented with the placement of the "Next" and "Previous" buttons, and explicitly surveyed web user's preference on the placement of the buttons (Bergstrom, et al., 2016). Their study revealed that their participants preferred the "Next" button to the right of the "Previous" (Bergstrom, et al., 2016). The findings also agreed that the location of the buttons made no difference in the completion time of the survey; however, participants eyes were drawn to the left button sooner than the right button when searching for the "Next" button (Bergstrom, et al., 2016). Evidently, the web users' actions contradicted their preference for the "Next" button on the right; however, their eyes were first drawn to the left when searching for the button. Since other studies claimed that respondents were more likely to unintentionally use the "Previous" button when it was to the right of the "Next" button, it may be better to follow the web users' preferred format.

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary of the recommended design considerations is presented in Table 2.

TABLE 2 – USER INTERFACE DESIGN CONSIDERATIONS RELATED TO PAGED WEB SURVEY NAVIGATION

Design Considerations	Reasoning
The design of the "Previous" button should be equally or less prominent than the "Next" button.	• When the "Previous" button is more prominently displayed, respondents often accidently click on the button which lengthened their completion times. • The "Next" button is more frequently used than the "Previous" button.
The "Next" button should be placed to the right of the "Previous" button.	• This is format is preferred by web users. • Respondents were more likely to unintentionally use the "Previous" button when it was to the right of the "Next" button

2.4 Colour

Marketing and psychology research has shown that colour can influence consumer's thoughts, emotions, and behaviour. Therefore, designers should take care in their colour selection so they can evoke the right emotions and keep respondents motivated throughout the survey. Table 3 summarizes qualitative research on the impact western culture has on the perception of the spectrum of colours.

TABLE 3 – THE INFLUENCE OF COLOR IN DESIGN

Colour	People's Unconscious Perception	Sources
Whites	• Sincerity, purity, cleanness • Simplicity, clarity, • Peace, happiness	(Wright, 1988) (Mahnke, 1996) (Fraser & Banks, 2004) (Clarke & Costall, 2007)(Labrecque & Milne, 2011)
Black and Greys	• Sophistication, glamour • Power, stateliness, dignity	(Odbert, et al., 1942) (Wright, 1988) (Mahnke, 1996) (Fraser & Banks, 2004) (Labrecque & Milne, 2011)
Yellows	• Cheerful, optimism, happiness • Sincerity, friendliness • Has a medium-long wavelength, thus has quality of arousal and excitement though less so than red	(Odbert, et al., 1942) (Murray & Deabler, 1957) (Wright, 1988) (Fraser & Banks, 2004) (Clarke & Costall, 2007) (Labrecque & Milne, 2011)
Reds	• Excitement, arousal, stimulating • Associated with characteristics of activity and strength • Research has consistently shown colours of longer wavelengths (i.e. reds, oranges, yellows) induces arousal and excitement	(Walters, et al., 1982) (Wright, 1988) (Fraser & Banks, 2004) (Clarke & Costall, 2007) (Labrecque & Milne, 2011)
Oranges	• Arousing, exciting but less so than red • Lively, energetic, extroverted, sociable	(Wexner, 1954) (Mahnke, 1996) (Labrecque & Milne, 2011)
Pinks	• Sincerity, nurturing, warm, soft • Feminine • Charming	(Mahnke, 1996) (Fraser & Banks, 2004) (Clarke & Costall, 2007) (Labrecque & Milne, 2011)
Purples	• Luxury, authenticity, quality • Feminine • Touching, nurturing	(Wright, 1988) (Mahnke, 1996) (Fraser & Banks, 2004) (Labrecque & Milne, 2011)
Blues	• Competence, intelligence, efficiency, logic • Communication, trust, duty, secure • Professionalism	(Murray & Deabler, 1957) (Wright, 1988) (Mahnke, 1996) (Fraser & Banks, 2004) (Labrecque & Milne, 2011)

Greens	• Nature, outdoors • Security	(Kaya & Epps, 2004) (Clarke & Costall, 2007) (Labrecque & Milne, 2011)
Browns	• Seriousness, reliability, support, protection • Nature, earthiness	(Wexner, 1954) (Murray & Deabler, 1957) (Wright, 1988) (Mahnke, 1996) (Fraser & Banks, 2004) (Clarke & Costall, 2007) (Labrecque & Milne, 2011)

There are very few technical and empirical studies on colour selection for website and web survey designs. One of the more scientific studies on web design colour selection was conducted by Pope and Baker (2005). Pope and Baker (2005) compared white, blue and pink backgrounds for a survey of college students. Their results revealed that background colour had no significant impact on response rates or drop-out rates of the survey; however, the survey completion times were slightly lowered for surveys with blue backgrounds compared to the pink and white backgrounds (Pope & Baker, 2005). The differences in completion times were larger for males than for females, which caused Pope and Baker to arrive at the conclusion that men are more negatively affected by pink backgrounds (Pope & Baker, 2005). Another similar study conducted by Baker and Couper (2007) involved the experimentation of the impact of white, blue and yellow background colours used on a consumer survey on energy use. They found that the drop-out rates were significantly higher for surveys with a yellow background (15%), while the surveys with a blue background had the lowest drop-out rates (10.8%) (Baker & Couper, 2007). However, there was no evidence in Baker and Couper's (2007) study showing that background colour had a significant impact on the actual completion time or the respondent's perceived completion times of the survey. As these studies, and several other user interface designers, suggest, there is a slight favour towards light neutral colours for backgrounds such as light blue (Tourangeau, et al., 2013; Baker & Couper, 2007; Pope & Baker, 2005).

In addition to carefully selecting colours based on user's conscious and unconscious preference, it is important to be aware of the reality of colour-blindness. According to the Colour Blindness Awareness Organization (2016), approximately 1 in 12 men and 1 in 200 women in the world have some variation of colour-blindness. Many of these colour-blindness disorders include difficulty distinguishing between blues and yellows, and reds and greens. A web cartographer, Muehlenhaus (2013) recommended the use of colour-blind tests/checks available on the web such as http://vischeck.com. Colour-blind tests/checks allow web designers to run a web page or image file through a website which simulates the page/image perceived by a person with colour-blindness (Vischeck, 2015). Below is a sample image illustrating colour-blindness simulation for a web-based map.

FIGURE 3 – COLOUR-BLINDNESS SIMULATION (EYEQUANT, 2013)

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary of the recommended design considerations is presented in Table 4.

TABLE 4 – USER INTERFACE DESIGN CONSIDERATIONS RELATED TO COLOUR SELECTION

Design Considerations	Reasoning
*Use light neutral **colours** for backgrounds such as light blue*	• Studies have shown that neutral background colours such as light blue can result in shorter completion times and lower drop-out rates. • Blue portrays professionalism, trust, and security which are all positive traits TTS would like to send to its respondents
*Avoid using the **colour** pink*	• Males appear to be more negatively affected by the colour. • Pink backgrounds have shown to cause increased completion times for males.
Conduct colour-blind tests/checks	• Survey design should accommodate for the colour-blind population.

2.5 Use of Images

One of the advantages of web surveys over telephone surveys are that images and diagrams can be used to aid respondents and to supplement the survey taking process. In web surveys, images may be

a decorative component of the user interface but can also serve as an integral part of a survey question.

Several studies in the literature have shown that images can sway survey responses and, much like colour, they can also influence respondents' emotions and thoughts. Couper and Tourangeau's (2004) investigated the impact of photographs accompanying six different survey questions, asking how often they performed an activity (ex. attending a sporting event during the past year). The pictures presented alongside the questions were strategically chosen such that the pictures represented either low or high-frequency exemplars of the question topic (i.e. a picture of sporting equipment by the sporting event question is a high-frequency exemplar) (Couper & Tourangeau, 2004). The results revealed an assimilation effect where respondents reported higher frequencies when higher frequency photographs were presented compared to lower frequency paragraphs (Couper & Tourangeau, 2004). Another study by Tourangeau et al. (2014) compared the effectiveness and the impact of visual versus verbal examples. It was found that pictures could narrow the respondent's interpretation of the question since pictures are more concrete and less general than verbal definitions (Tourangeau, et al., 2014). For example, in the study, the respondents were asked to report their fruit consumption, and a sample of the respondents was supplemented with a visual example such as a picture of an apple and others were provided with a verbal definition of fruits (Tourangeau, et al., 2014). The results showed that, on average, respondents reported that they consumed more fruit when they were provided with a verbal definition compared to just a picture of an apple (Tourangeau, et al., 2014).

Tourangeau, Couper, and Galešic (2005) also experimented with the placement of images. They conducted an eye-tracking study and found that pictures placed in the header are often ignored by survey respondents (Tourangeau, et al., 2005). This may be explained by the "banner blindness" phenomenon discussed earlier in Section 2.1. Note that the study by Shrestha and Lenz (2007), also discussed in Section 2.1, found that the "F-pattern" scanning pattern did not hold true when a user was browsing or searching through a picture-based webpage (i.e. blocks of pictures instead of text).

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary of the recommended design considerations is presented in Table 5.

TABLE 5 – USER INTERFACE DESIGN CONSIDERATIONS RELATED TO COLOUR SELECTION

Design Considerations	Reasoning
*When supplementing a survey question with **an** example, opt for written examples/definitions over visual examples.*	• Images have been shown to narrow respondent's interpretation of the survey question.
Be aware that images placed in the header are often ignored.	• "Banner Blindness" phenomenon

2.6 Typography

Selecting the typeface and font sizes for a web survey may appear like a trivial task; however, the typography used in user interface design should be done with care, since research has shown that

typeface can convey meaning and affect how questions are interpreted (Childers & Jass, 2002; Micheal S. McCarthy, 2002; Novemsky, et al., 2007). Most of the literature on typeface selection and design is qualitative, such as the study by Ling and Schaik (2006) who compared a serif font (i.e. Times) with a san serif font (i.e. Arial) on a web page. The participants of the study preferred Arial; however, the results showed no significant difference in user behaviour between the two typefaces (Ling & Schaik, 2006). A similar study by Bernard et al. (2003) arrived at the same conclusion that there are no differences in performance across the use of different typefaces. On the other hand, Schmidt et al.'s (2009) findings contradicted this with their study revealing that their participants performed tasks quicker with serif fonts. However, like Ling and Schaik's study, Schmidt et al. (2009) found that web users preferred san serifs fonts. The popularity of san serif fonts among web-users may be explained by their widespread use, due to the design guidelines set by the World Wide Web Consortium (W3C), an international community that develops web standards (W3C, 2016).

There exist very few technical studies on font sizes used in web design since the interpretation of legible, comfortable reading size is subjective. A study by Bernard et al. (2003) compared 12- and 14-point size fonts, and found that the 14-point size font was more legible for users and produced faster reading speeds. However, this study did not provide a fair indication of the font size that should be used in survey designs, as the study was limited to texts viewed on a PC and did not explore, in-depth, the large range of available font sizes.

The line length used in the display of text on a webpage can influence the behaviour of web users. Ling and Schaik (2006) conducted a study that found that web-users preferred shorter line lengths of text. However, they found that web-users performed faster with longer line lengths when conducting visual search/browsing tasks since it allowed for faster scanning and fewer lines need to be scanned (Ling & Schaik, 2006). Therefore, Ling and Schaik (2006) suggested that longer line lengths should be used for information that is designed to be scanned quickly, and shorter line lengths should be used when the text should be read more thoroughly. Earlier studies by Dyson et al. also supported this statement (Dyson & Kipping, 1998; Dyson & Haselgrove, 2001).

Clarity and readability in a user interface design are important in creating user-friendly designs. Research has suggested that clarity and readability can be improved by making the webpage appear less crowded. One of the ways to achieve this has been by using longer line lengths, which caused the illusion of fewer text since the text fits on fewer lines. According to Galitz (2002), web-users have been found to be 20% more productive on less-crowded web-pages, completing transactions in 25% less time and with 25% less error than those with a more crowded screen.

However, there are times where de-crowding a screen of text is unachievable without compromising the information on the webpage. A solution to drawing the web user's attention to key information within a block of text is to emphasize the selective text with boldface, underlines, capitalization, and/or use of colour. Once again, the majority of the literature on selective emphasis in the text are qualitative and descriptive. Nielsen (2004) states that the colour blue and/or underlining should not be used to emphasize text, for they are conventionally used as a linkable text in web design. Tourangeau et al. (2013), suggested that emphasizing a long body of text with capitalization was not as effective compared to the capitalization of smaller text segments (Tourangeau, et al., 2013). They also suggested that italicizing text can draw emphasis due to its contrast with regular text; however, it should be avoided since it is less readable (Tourangeau, et al., 2013). Tourangeau et al. (2013)

pointed out that italics are conventionally used for instructions in web surveys, and that italics give respondents the notion that the question text is not as important and can be ignored. Lynch and Horton (2001) warned that too much emphasis can be counter-productive. They also suggested that if the regular text is used for questions, bolding is the most effective for emphasis and that capitalization is the most effective in all non-regular text (i.e. headings) (Lynch & Horton, 2001). Finally, Lynch and Horton (2001) advised that the use of text styles for emphasis should be consistent throughout the web survey so that respondents could learn the meaning associated with each text element through repetition.

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary is presented in Table 6.

TABLE 6 – USER INTERFACE DESIGN CONSIDERATIONS RELATED TO TYPOGRAPHY

Design Considerations	Reasoning
Use san serif fonts, such as Arial, over serif fonts (ex. Times New Roman)	• Web-users have shown a preference for san serif fonts. • There is a widespread use of sans serif fonts due to the design guidelines set by the World Wide Web Consortium (W3C).
Longer line lengths should be used for information that is designed to be scanned quickly, and shorter line lengths should be used when the text should be read more thoroughly	• Web-users perform faster with longer line lengths when conducting visual search/browsing tasks since it allowed for faster scanning and fewer lines need to be scanned. • Web-users have shown a preference for shorter line lengths for they are easier to read.
Avoid emphasizing text using the colour blue and underlining	• Blue text with and without underlining is conventionally used for linkable text
Opt for boldface over capitalization when emphasizing text in large text blocks. Use capitalization to emphasize text in headings.	• Emphasizing a long body of text with capitalization is not as effective compared to the capitalization of smaller text segments
Avoid italicizing text unless it is for unimportant information that can be ignored.	• Italicized text is less legible. • Italics are conventionally used for instructions in web surveys and it give respondents the notion that the question text is not as important and can be ignored

2.7 Web Accessibility Guidelines

When developing a public website, it is important to consider the diverse needs of the users, especially the needs of people with disabilities. There exist a large range of disabilities that can impair a person's accessibility to a website such as blindness, deafness, hearing loss, learning disabilities, cognitive limitations, limited movement, speech disabilities, photosensitivity and combinations of theses (Ontario, 2017). Therefore, to ensure web-content accessibility to the diverse public, the Government of Ontario requires new public websites by law to meet the Web Content Accessibility Guidelines (WCAG 2.0) (Ontario, 2017).

The WCAG provides three levels of web-accessibility guidelines (A, AA, and AAA) where each level enforces stricter accessibility standards than its predecessor. In many cases, to achieve a higher level of accessibility, the standards of the previous level must first be achieved. Since January 2014, the Government of Ontario requires all new and significantly updated websites to comply with WCAG 2.0's level A standards. By January 2021, all public websites posted after January 2012 must meet WCAG 2.0 level AA standards, except for sections 1.2.3. and 1.2.5 (Ontario, 2017). The design of the new TTS web-survey is required to meet at least the level A standard, and will soon be required to meet the level AA standards. Therefore, Table 7 provides a summary of the guidelines to be followed by the new TTS web-survey. As organized in the table, the WCAG guidelines can be categorized under the following four main categories which are synonymous to the guidelines' accessibility goals: perceivable, operable, understandable, and robust (WCAG, 2016).

The current state of the TTS web-survey design complies with most, if not all the level A and level AA accessibility standards outlined in Table 7. The survey is mainly comprised of electronic text, and does not have any time-based or multi-media where captions and alternative contents needs to be provided. However, the Google Map API used in location-based and trip-based questions is not necessarily keyboard accessible as required by WCAG 2.0 Section 2.1. For example, a user would not be able to drag and drop pins on the map or adjust waypoints along routes using only their keyboard. Fortunately, the Ontario Government recognizes that there are cases where it may not be possible to post content that complies with WCAG 2.0, such as online maps and complex diagrams (Ontario, 2017). In these cases, the Ontario E-laws states that the content may still be posted, but accessible formats should be provided upon request. Therefore, alternative formats such as telephone surveys should be provided for the visually disabled upon request.

TABLE 7 - SUMMARY OF WEB CONTENT ACCESSIBILITY GUIDELINES (WCAG) 2.0 FOR LEVEL 'A' AND LEVEL 'AA' STANDARDS (WCAG, 2016)

PERCEIVABLE		
Guideline	***Level***	***Description***
1.1: Provide text alternatives for non-text content	***A***	• *Provide text alternatives for any non-text content so that if can be changed into other forms people need (large print, braille, speech, symbols, or simpler language).* • *Text should be electronic text, not an image of text.*
1.2: Provide alternatives for time-based media	***A***	• *Provide captions and other alternatives for multimedia.* • *Provide a document presenting the same information as the prerecorded audio-only content.*
	AA	• *Provide a descriptive audio track and/or other alternatives for prerecorded video-only content.* • *Captions are provided for all prerecorded audio content and all live audio content in synchronized media.*
1.3: Adaptable content - Present content that can be presented in different ways	***A***	• *Information, structure, and relationships conveyed through presentation can be programmatically determined or are available in text.* • *Information and relationships that are implied by visual or auditory formatting are preserved when the presentation format changes.*

without losing information or structure		• *When the sequence in which the content is presented affects its meaning, a correct reading sequence can be programmatically determined.* • *Instructions provided for understanding and operating content do not rely solely on sensory characteristics of components (shape, size, visual location, orientation, or sound).*
1.4: Distinguishable – Make it easier for users to see and hear content including separating foreground from background	***A***	• *Color is not used as the only visual means of conveying information, indicating an action, prompting a response, or distinguishing a visual element* • *If an audio plays automatically for more than three seconds, there must be a mechanism that pauses/stops the audio, or allows audio volume control independent from the overall system volume level.*
	AA	• *The visual presentation of text and images of text has a contrast radio of at least 4.5:1. Exceptions include large text, incidental purely decorative text or images, and logotypes.* • *Except for captions and images of text, text can be resized without assistive technology up to 200% without loss of content or functionality.* • *If technologies being used can achieve the visual presentation, text is used to convey information rather than images of text.*
OPERABLE		
Guideline	***Level***	***Description***
2.1: Make all functionality available from a keyboard	***A***	• *All functionality of the content is operable through a keyboard interface without requiring specific timings for individual strokes.* • *Ensure that the content does not "trap" keyboard focus within subsection of content on a webpage. This is a common problem when multiple formats are combined within a page and rendered using plug-ins or embedded applications.*
2.2: Provide users enough time to read and use the content	***A***	• *For each time limit that is set by the content, the user is allowed to turn off, adjust, extend the time limit.* • *Unless essential, for any moving, blinking, or scrolling information that starts automatically, lasts more than five seconds, and is presented in parallel with other content, a mechanism should be provided to allow users to pause, stop, or hide it.*
2.3: Do not design content in a way that tis known to cause seizures	***A***	• *Web pages do not contain anything that flashes more than three time in any one second period, or the flash is below the general flash and red flash threshold.*
2.4: Provide ways to help users navigate, find content, and determine where they are	***A***	• *A mechanism is available to bypass blocks of content that are repeated on multiple web pages.* • *Web pages have titles that describe topic or purpose.* • *If the navigation sequence of a webpage affects meaning or operation, focusable components receive focus in an order that preserves the meaning and compatibility.* • *Purpose of each link can be determined from the link text alone, except where the purpose of the link would be ambiguous to users in general.*

	AA	• *Provide more than one way to locate a web page within a set of webpages, except when the webpage is a result of a process.* • *Headings and labels describe topic or purpose.* • *Any keyboard operable user interface has a mode of operation where the keyboard focus indicator is visible.*

UNDERSTANDABLE

Guideline	**Level**	**Description**
3.1: Make content readable and understandable	**A**	• *The default language of each page can be programmatically determined.*
	AA	• *The language of each passage or phrase in the content can be programmatically determined except for proper names, technical terms, words or indeterminable language, and words or phrases that have become part of the vernacular of the immediate surrounding text.*
3.2: Make web pages appear and operate in predictable ways	**A**	• *When any component receives focus, it does not initiate a change of context.* • *Changing the setting of any user interface component does not automatically cause a change in context unless the user has been advised of the behavior before using the component.*
	AA	• *Components that have the same functionality within a set of web pages are identified consistently.*
3.3: Help users avoid and correct mistakes	**A**	• *If an input error is automatically detected, the item that is an error is identified and the error is described to the user in text.* • *Labels or instructions are provided when content requires user input.*
	AA	• *If an input error is automatically detected and suggestions for correction are known, then the suggestions are provided to the user, unless it would jeopardize the security or purpose of the content.* • *For web pages that cause legal commitments or financial transactions for the user to occur, that modify or delete user-controllable data in data storage systems, or that submit user test responses, at least the following is true: 1) submissions are reversible; 2) Data is checked for input errors and user is provided an opportunity to correct them; 3) mechanism is available for reviewing, confirming, and correcting information before finalizing the submission.*

ROBUST

Guideline	**Level**	**Description**
4.1: Maximize compatibility with current and future user agents, including assistive technologies	**A**	• *In content implemented using markup languages, elements have complete start and end tags, elements are nested according to their specifications, elements do not contain duplicate attributes, and any IDs are unique, except where the specification allow these features.* • *For all user interface components, the names and role can be programmatically determined; states, properties, and values that can be set*

		by the users can be programmatically set; and notification of changes to these items is available to user agents, including assistive technologies.

3 DESIGN OF WEB SURVEY QUESTIONS

A key process in the design of web surveys is planning the layout and presentation of the questions. The layout and presentation of a survey can influence the respondent's interpretation of the questions. Thus, it is important to design for clear, comprehensive survey questions to avoid misinterpretations and unnecessary burden on the respondents.

This section of the report presents a literature review of the design of web survey questions. The purpose of the literature review is to identify elements in web survey question design that will improve the usability of the future TTS web survey. The literature review will also aid the survey designers with the selection of question types for the various TTS household questions.

3.1 Wording and Structure of Survey Questions

Empirical studies on survey question wording and design were not found. Much of the literature offered general guidelines set by survey design experts. These guidelines are applicable to all types of questionnaires, including web surveys. Many of these guidelines are intuitive; however, they are often overlooked by surveyors. Table 8 below summarizes some common guidelines on the wording and structure of survey questions that have been enforced by various experts in the field. It is important to consider these guidelines in the design of the TTS web survey, as they have the potential to ease the survey-taking the process for respondents.

TABLE 8 – GENERAL GUIDELINES FOR THE DESIGN OF SURVEY QUESTIONS

Category	Guidelines	Sources
Clarity of Wording	• Keep instructions and text simple and unambiguous • Avoid the following: - Jargon - Complex words and technical terms - Shorthand (i.e. contractions, abbreviations, symbols) - Framing questions in the negative - Using double negatives - Passive voice - Words or phrases with strong point of view that may sway responses	(UWSC, 2010) (Galitz, 2002) (Krug, 2006) (Lee-Gosselin Associates Limited, 2012) (Tourangeau, et al., 2013)
Structure	• Should not combine two questions into one • Order the questions in a logical manner • Response categories should be exhaustive and mutually exclusive • Group questions that are similar in topic, and group questions within that topic with similar response options/formats. Repetition improves the respondent's time to complete the survey and their ability to understand the survey structure. • Maintain a parallel structure for all questions: - Use same words or phrase to refer to the same concepts - Position repeated words and clauses in the same place and in the same order	(UWSC, 2010) (Galitz, 2002) (Krug, 2006) (Lee-Gosselin Associates Limited, 2012) (Tourangeau, et al., 2013)

3.2 Radio Buttons and Multi-Select Checkboxes

The radio button is the most commonly used response format in surveys. It is used for single response questions and is conventionally displayed as a hollow circle alongside each response option. When a response option is selected, the hollow circle is filled with a solid circle. Figure 4 is an example of a radio button question. Checkbox questions are similar to radio button questions except they are used for multi-response questions. The conventional design is a hollow square beside each response option, and when a response is selected, a checkmark is displayed within the hollow square. Figure 5 is an example of a multi-select checkbox question. Note that it is important to follow the conventional designs of survey questions, for they can lessen or eliminate the learning curve for the respondents.

FIGURE 4 – CONVENTIONAL DESIGN OF A RADIO BUTTON QUESTION

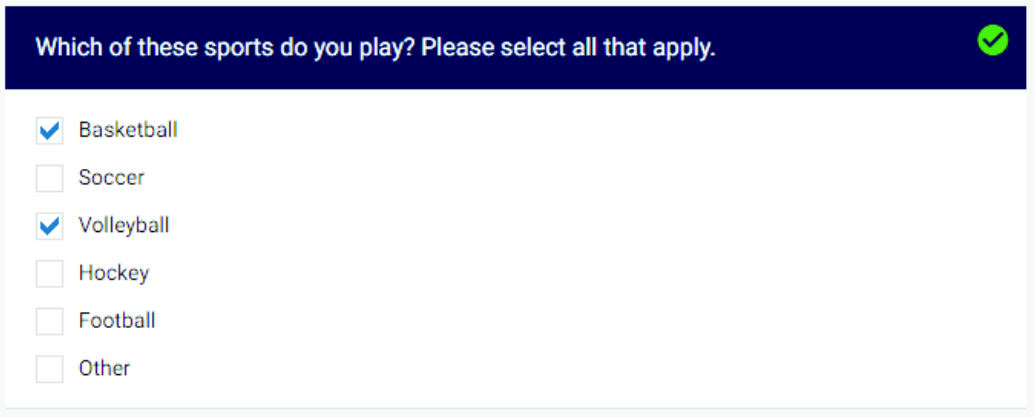

FIGURE *5* - CONVENTIONAL DESIGN OF A MULTI-SELECT CHECKBOX QUESTION

Many studies have found that in vertically displayed response options for radio button questions and checkbox questions, respondents tend to pay greater attention to the earlier options compared to options near the bottom. For example, Couper et al. (2004) studied how the relative visibility of responses affects the behaviour of the respondents. The study compared three question formats: radio buttons which display all response options, a drop-down box that hides responses until the user clicks the drop-down arrow, and a drop-down box that initially displays five options and hides the other five (Couper, et al., 2004). Couper et al. (2004) also varied the order of response options in the surveys to avoid bias results. The study revealed that response options visible from the offset (i.e. radio button format, and the first five visible dropdown options) were 27% more likely to be selected compared to the initially hidden responses (Couper, et al., 2004). Galesic et al. (2009) conducted a

similar study for radio button and checkbox questions, instead of tracking the eye movements of the participants. The eye-tracking data revealed that respondents fixated longer on the options near the beginning compared to the end (Galesic, et al., 2009). Their findings suggested the visibility of response options arranged vertically decreases with the response order (Galesic, et al., 2009).

The display of many survey questions, including questions in radio buttons and multi-select checkboxes, differ on mobile phones compared to personal computers due to the difference in their screen sizes. Often on smaller screens, response options appear closer together. As proven in various studies, closer responses may cause respondents to accidently select the wrong response options (Tourangeau, et al., 2013). Peytchev and Hill's (2010) mobile-survey experiment observed that many of their respondents were willing to navigate to response options initially hidden; however, 23% of the respondents explicitly indicated that they noticed additional response options, but complained it took too much effort to unveil them. Stapleton (2013) compared the response behaviour for horizontal radio button question on a mobile device and a personal computer. Respondents were more likely to select the first option (i.e. left most option) when completing the survey on a mobile device (84.7%) than on a computer (78.8%) (Stapleton, 2013). Therefore, Stapleton (2013) concluded that the visibility of responses has a greater impact on the response selection of mobile respondents versus computer respondents.

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary of the recommended ones is presented in Table 9.

TABLE 9 – DESIGN CONSIDERATIONS FOR RADIO BUTTONS AND MULTI-SELECT CHECKBOX QUESTIONS

Design Considerations	Reasoning
For vertical display of response options, display most likely response options at the top of the list.	• Visibility of response options displayed vertically decreases with the response order.
For horizontal display of response options, display most likely response options at the left of the list.	• Respondents were more likely to select the first option (i.e. left most option). This has a more pronounced effect on mobile users.

3.3 Single Select Dropdown Menus

Dropdown menus are typically used for lengthy lists of response options to avoid making the survey appear unnecessarily crowded. For extremely lengthy lists, a subset of the responses is displayed when the respondents click on the drop-down arrow and the remaining options are displayed as the respondent scrolls down. An example of a drop-down menu question is displayed as Figure 6. Conventionally, drop down menus are used for single response questions; however, there exist a few web survey designs that allow for multi-response selection.

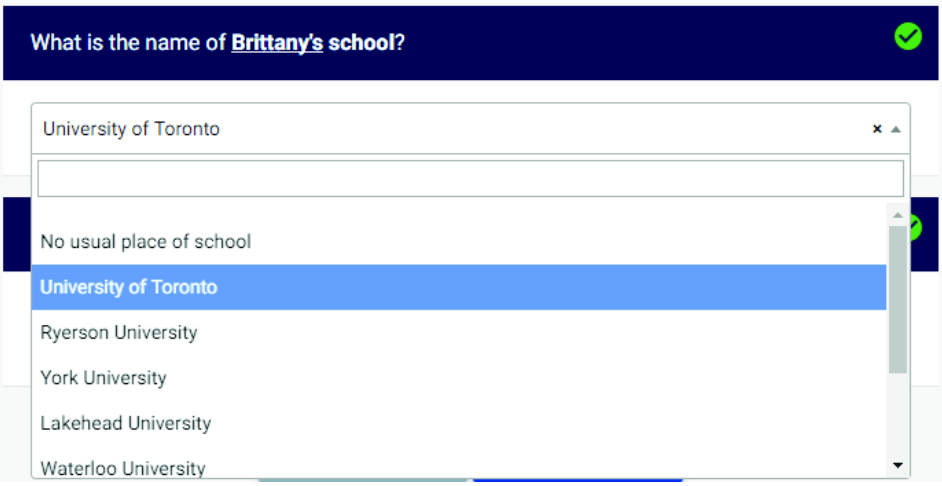

FIGURE 6 - CONVENTIONAL DESIGN OF A SINGLE-SELECT DROPDOWN MENU QUESTION

The concept of visibility of responses discussed in Section 3.3 is also applicable to drop-down menus. In summary, respondents are more likely to select options nearing the beginning of the list since they are more readily visible. However, for short lists of response options, Heerwegh and Looseveldt (2002) and Healy (Healey, 2007) found that survey completion rates did not differ between radio buttons and drop-down menus. Rationally speaking, it would appear that a series of questions in dropdown menu format compared to radio buttons would take slightly more time and effort by the respondent because of the extra click required to expand the drop-down menu. Unfortunately, guidelines on when to choose radio buttons over drop down menus based on the number of response options were not found during this literature review.

An investigation of the 2009 Canadian Census results by Labrasseur et al. (2010), revealed an issue with dropdown menus that potentially resulted in some inconsistencies in responses. They found that 76% of respondents used a scroll mouse in the census' web survey and, thus, were susceptible to accidently changing their answers in the drop-down menu as they scrolled (Lebrasseur, et al., 2010).

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary of the recommended ones is presented in Table 10.

TABLE 10 – DESIGN CONSIDERATIONS FOR SINGLE SELECT DROPDOWN MENU QUESTIONS

Design Considerations	Reasoning
Display most likely response options at the top of the list.	• Visibility of response options displayed vertically decreases with the response order.
For short to medium sized response options list, opt for the radio button format.	• Response options displayed in radio button format is more visible compared to the drop-down menu format. • The dropdown menu requires an additional mouse click.

	• Respondents are more susceptible to accidently changing their answers in the drop-down menu as they scroll down.

3.4 Text Boxes

Text boxes are designed for open-ended questions for text, dates and other numerical responses. An example of a text box formatted question requiring a numerical response is shown in Figure 7 below.

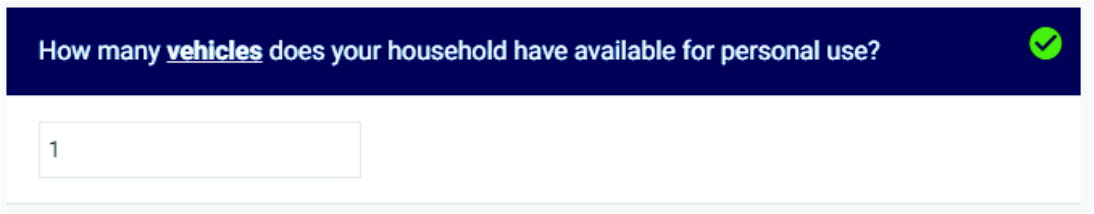

FIGURE 7 - CONVENTIONAL DESIGN OF A SINGLE-SELECT DROPDOWN MENU QUESTION

Text boxes are conventionally displayed as a type-able area bounded by a border. Due to its simplistic design, the key design consideration for textboxes is their size. Intuitively, a larger box implies the requirement of a longer, more detailed answer. Couper, Traugott, and Lamias (2001) investigated the impact that text box sizing has on responses, finding that longer boxes produced more invalid responses than short boxes. This was because text boxes allowed for respondents to avoid answering questions, and often did not prevent out-of-range or irrelevant responses. The exception was where the textbox was formatted for dates/numbers and was subjected to conditions. Also, text boxes required more time and effort by respondents compared to a click of a radio button. The response burden associated with text boxes often deterred respondents from completing the question.

Textboxes for numerical responses often have more design elements than just a simple border. The design includes graphical hints to the respondent on the required response format or structure. For example, Couper et al. (2011) conducted an experiment on an open-response question asking for a dollar amount. They reported that respondents were most likely to conform to the desired format (i.e. amount rounded to the nearest whole dollar) when the textbox displayed a dollar sign to the left, and a decimal followed by two zeros on the right (Couper, et al., 2011). Fuchs (2009) had similar findings, recommending that open-ended questions should include a label, and not just a blank box (i.e. "_____ no. of students" versus "_____"), to promote exact and appropriate answers.

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary is presented in Table 11.

TABLE 11 – DESIGN CONSIDERATIONS FOR TEXT BOX QUESTIONS

Design Considerations	Reasoning
When deciding between a closed question format (ex. Radio button) and a textbox, opt for the closed question format.	• Text boxes require more time and effort by respondents compared to say a click of a radio button. • Text boxes often deter respondents from completing the question. • Closed question formats prevent an invalid and out-of-range response.
The size of the textbox should indicate the length and details of an appropriate response.	• Larger box implies the requirement of a longer, more detailed answer.
*Use graphical hints and labels beside textboxes (ex. dollar sign to the left of the text box and a decimal place followed by two zeros the right of the **text box** indicates to the respondent to input a monetary amount rounded to the nearest dollar).*	• Promotes exact and appropriate answers.

3.5 Sliders and Likert Scales

Sliders are typically continuous scales with a specified upper and lower range. It is an alternative method for collecting bounded numerical responses. It provides a respondent with an obvious visualization of the question's upper and lower limits. Respondents can indicate a number by dragging the marker along the scale. Figure 8 is an example of a slider question. On the other hand, Likert scales are discrete scales. They display a discrete number of response options which can be numeric, a descriptive rating system, or any group of relatable options with the hierarchical/ chronological order. An example of a Likert scale question is presented as Figure 9.

FIGURE 8 - CONVENTIONAL DESIGN OF A SLIDER QUESTION

FIGURE 9 - CONVENTIONAL DESIGN OF A LIKERT SCALE QUESTION

Although both scale-type formats appear similar, an experiment by Couper et al. (2006) revealed that the respondents' behaviour and responses slightly varied between the two formats. Couper et al. (2006) found that sliders produced higher survey drop-out rates and question non-response than Likert scales, as well as longer response times. However, their web survey had issues with the display of the sliders. They believed that it was attributed to the compatibility of the installed/enabled Java (Couper, et al., 2006). The Student Move TO web survey also experienced issues with the display of their slider question. Even with removing the download issue from their results, the respondents appeared to still favour the Likert scales over the sliders (Couper, et al., 2006). Couper et al. (2006) noted that the distribution of responses did not differ significantly between the two formats; however, respondents selected the mid-point response option more often in the Likert format.

The position of the mid-point response option on scale-type questions also influences the response results. Tourangeau et al. (2004) demonstrated this by analyzing the distribution of responses when the conceptual mid-point option did not visually appear in the midpoint of the scale (i.e. scale options are not evenly spaced) compared to when conceptual mid-point and visual mid-point coincided. Their results found that respondents relied on the visual midpoint as their reference compared to the conceptual midpoint (Tourangeau, et al., 2004). For example, when respondents were presented with a probability scale and the conceptual mid-point (i.e. even chance, 50%) was placed further to the left of visual centre of the scale, respondents were more likely to select an option on the lower probability end of the scale than when they were presented with an evenly distributed scale (Tourangeau, et al., 2004). The results from Tourangeau et al.'s (2004) study suggests that surveyors should design for evenly distributed response options for scale-type questions and have the conceptual and visual midpoints coincide on the scale to avoid skewed responses.

The choice of numbers for a numerical scale may introduce bias in the response results. Schwarz et al. (1991) compared the response results for a bipolar scale question using negative numbers (e.g. -5 to +5) and a positive number scale (e.g. 0 to 10). When the scale ranged from 0 (i.e. "not at all successful") to 10 (i.e. "extremely successful"), 34% of respondents selected values between 0 and 5 (Schwarz, et al., 1991). However, when respondents were presented with a bipolar scale ranging from -5 (i.e. "not at all successful") to 5 (i.e. "extremely successful"), only 13% of respondents selected values between -5 and 0 (Schwarz, et al., 1991). Schwarz et al. (1991) findings revealed that respondents using the bipolar scale tended to select positive numbers, which significantly increased the positivity bias of the results. Schwarz et al. (1991) also suggested that different interpretations arose when respondents attempted to associate the numeric scale values with the meaning of the scale labels. Therefore, using numeric scales/sliders in surveys should be constructed with care as it may introduce bias and compromise the quality of data collected.

To avoid bias that may arise from the use of numeric scales, many survey designers have recommended the use of Likert scales with verbal labels. Again, care should be taken when selecting verbal labels for response options for they too can sway responses. Tourangeau et al. (2004) found that the order of verbal labels arranged horizontally (i.e. gradient from negative left to positive-right, versus gradient from positive left to negative-right) had no significant effect on the response time or distribution of responses. However, they also found that this did not hold true for Likert scales arranged vertically. Meier and Robinson (2004) revealed similar findings. Respondents were found to identify positive labels (i.e. brave, loyal, hero, etc.) more quickly when they were positioned at the top of the scale compared to being placed at the bottom of the scale. Conversely, respondents were

also quicker to identify negative labels (i.e. bitter, clumsy, crime) when they were placed at the bottom of the scale (Meier & Robinson, 2004). In summary, these studies revealed that respondents are more accustomed to a positive (top) to negative (bottom) gradient for a vertical Likert scale.

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary of the recommended ones is presented in Table 12.

TABLE 12 – DESIGN CONSIDERATIONS FOR SLIDERS AND LIKERT SCALE QUESTIONS

Design Considerations	Reasoning
When deciding to use sliders versus Likert scales, opt for a Likert scale design.	• Sliders compared to Likert scales produce higher survey drop-out rates and question non-response, as well as longer response times. • Some web surveys had issues with the display of the sliders. • Respondents prefer Likert scale questions to sliders.
*In the use of numerical scales, use only a positive range of numbers. Avoid **bipolar** scales.*	• Respondents using the bipolar scale tended to select positive numbers, which significantly increased the positivity bias of the results.
Design for evenly distributed response options for scale-type questions and have the conceptual and visual midpoints coincide on the scale	• Unevenly distributed response options have shown to skew survey results. • Respondents rely on the visual midpoint as their reference.
For vertical Likert scales, adopt a positive (top) to negative (bottom) gradient.	• Respondents have been found to be accustomed to this format
Take caution in selecting verbal labels for Likert scales.	• Verbal labels can sway survey results.

3.6 Grid/Matrix Questions

Multiple questions that share the same response options can be displayed in a matrix format. Formatting the shared responses in the grid tends to use less screen space and eliminates repetitive text, thereby reducing page clutter. Figure 10 is an example of a survey question displayed in a matrix format.

The grid format also eliminates the need for the respondents to repetitively read the shared response options. This leads many web survey designers to believe that the matrix format can shorten survey completion times. Couper, Traugott, and Lamias (2001) tested this theory on a web survey containing five knowledge measure questions and eleven attitude measure questions. Each type of question was presented in two different versions; the five knowledge questions were presented as singular questions on five separate pages or a matrix question on a single page, and the eleven knowledge questions were presented as singular questions on eleven separate pages or as three matrix questions on three separate pages (Couper, et al., 2001). Couper, Traugott, and Lamias' (2001) results showed that respondents completed the matrix version of the questions significantly quicker compared to the

singular question version (i.e. the average completion time for sixteen matrix formatted question was 168 seconds, compared to 194 seconds for the singular formatted version). An experiment by Bell, Mangione, and Kahn (2001) reported similar findings. They reported a shorter completion time for the matrix version of their health questionnaire compared to the single question per page version (i.e. 5.22 versus 5.07 minutes) (Bell, et al., 2001). Later studies, such as those by Tourangeau et al. (2004), and Toepoel, Das and van Soest (2008) arrived at the same conclusion that the matrix format can shorten completion times.

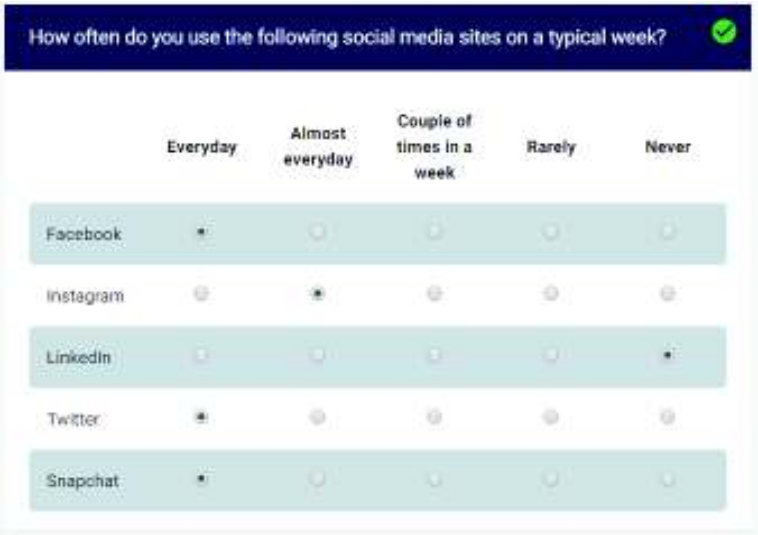

FIGURE 10 - CONVENTIONAL DESIGN OF A MATRIX QUESTION

The lower completion time for a matrix question may be attributed to the fact that the grid format makes it easier for respondents to quickly complete the survey by selecting the same responses with minimal scrolling and mouse movement. Couper, Traugott, and Lamias' (2001), and Toepoel, Das and van Soest (2008) evaluated the correlation between the responses to subsequent questions when displayed in a matrix versus singular formats; they found that the correlation of responses was slightly higher for the matrix version. Furthermore, Couper, Traugott, and Lamias (2001) claimed that the rates of explicit non-responses (i.e. selecting the "Don't know" or the "Not applicable" options) were also higher for the matrix version. Tourangeau et al. (2004) compared three different versions of their eight-question web survey: 1) all questions formatted in a matrix on a single page, 2) two matrices on a single page containing four questions each, 3) and a single question per page. Their survey results showed a significant linear trend with increasing correlation of responses for subsequent questions as the number of grouped questions in matrices increased (Tourangeau, et al., 2004). Furthermore, Tourangeau et al. (2004) suggested that respondents tended to repetitively choose the same response options for all grouped questions in a matrix format, for they noticed that there were fewer discrepencies between answers for versions 1 and 2 compared to version 3 of their web survey. Bell, Mangione and Kahn (2001), on the other hand, found no significant difference in correlation of subsequent question responses between the two survey versions. It is important to note that the inconsistencies of correlation reports may be attributed to the design of the web surveys, variation in the sample population, differences in questionnaire topics or the number of response options. However, the majority of these studies suggested that formatting similar questions in a matrix may

compromise the reliability of the responses. Unfortunately, none of the studies recommend an acceptable range of questions to be grouped in a matrix.

Based on the findings of the literature review, several design considerations for the future TTS web survey were devised. A summary of the recommended ones is presented in Table 13.

TABLE 13 – DESIGN CONSIDERATIONS FOR MATRIX QUESTIONS

Design Considerations	Reasoning
Use a matrix format to combine simple questions sharing the same response options.	• Matrix format can shorten completion times.
Avoid using the matrix format for questions that require more thought by the respondent.	• Respondents tend to satisfice through the survey by repetitively selecting the same response options for all grouped questions. • Matrix questions are shown to produce higher rates of explicit non-responses (i.e. selecting the "Don't know" or the "Not applicable" options).

4 DESIGNING FOR ACCURATE AND COMPLETE RESPONSES

Large surveys, such as the TTS, are costly to undertake, thus it is important to maximize a survey's utility through careful planning, design, and execution. In order to ensure cost-effectiveness of a survey, the data collected should be as accurate and precise as possible. Applying the user interface and survey question design elements discussed earlier in this report can help achieve accurate and complete responses. There are further steps in survey design, beyond the aesthetics and formatting of the survey questions and user interface, that can also be taken to motivate respondents.

This section of the report attempts to identify web survey elements proven in the literature that can improve the efficiency and success of a web survey. The purpose of the literature review was to identify and understand key web survey elements that directly and/or indirectly influence the accuracy/completeness of survey responses, and how they can be properly applied to the future TTS web survey.

4.1 Introducing the Survey

It is understandable that the topic and sponsor of a survey both have considerable weight on a respondent's decision to participate in a survey. Various studies in the literature have found that respondents were more likely bear the time and burden of a survey if they had an interest in the survey topic and/or the results of the survey (Sills & Song, 2002; Alsnih, 2006; Tourangeau, et al., 2013). However, the findings of an experiment conducted by Tourangeau et al. (2009) contradicted this finding. Tourangeau et al. (2009) asked members to take part in a web survey based on the description of the survey provided to them. The descriptions were systematically varied to include different combinations of topics and sponsors (ex. "The National Coalition of Gun Owners", "The National Coalition for Victims of Gun Violence", and "The National Center for the Study of Crime") (Tourangeau, et al., 2009). Their study reported that the member's interest in the topic did not significantly impact the response rate of the survey (Tourangeau, et al., 2009). Other than this study, there has been very little empirical research on the relationship between survey topic/sponsor and response rates. However, some studies have argued that surveys sponsored by government agencies or an academic researcher achieve higher response rates than a survey sponsored by a commercial firm (Tourangeau, et al., 2013).

Many web surveys are distributed through e-mail invitations. In e-mail invitations, the first impression a respondent will have of the survey is based on the subject line of the e-mail. This has been recognized by various web survey designers, and there exists various research about the influence of e-mail subject lines on the response rate of surveys. For example, Kent and Brandel (2003) experimented with prize-subject lines, such as "Win a weekend for two", to motivate members of a customer loyalty program to complete an online survey. They reported that the prize subject line produced a lower response rate (52%) compared to a subject line that explicitly stated the e-mail was about a survey (68%) (Kent, & Brandel, 2003). Trouteaud (2004) experimented with plea-subject lines (i.e. "Please help [Company Name] with your advice and opinions") and offer subject lines (i.e. "Share your advice and opinions now with [Company Name]"). They found that the "plea" subject line yielded 5% higher response rate compared to the surveys with offer subject lines (Trouteaud, 2004).

Other studies have investigated the impact that personalization of e-mails has on survey response rates. Joinson (2007) has performed several studies in this field, finding that personalization was only effective when the sender of the e-mail/survey sponsor was of high status and that personalization tended to decrease the respondent's sense of anonymity, deterring them from disclosing personal information in the survey (Joinson, et al., 2007; Joinson, et al., 2010). Joinson (2007) also found that survey sponsors with higher status, such as a government agency or academic researcher, yielded higher response rates; this statement was also supported by a study by Gueguen and Jacob (Guéguen & Jacob, 2004). Heerwegh et al. (2005) experimented with personal salutations (i.e. "Dear [First name and Last name]") and impersonal salutations (i.e. "Dear student) for an online survey distributed to college students. They found that students who received the personal and impersonal salutations had response rates of 64.3% and 54.5%, respectively. However, Pearson and Levine (2003) disagreed with these findings for their study on Stanford University alumni, revealing that personalization offered a small but insignificant influence on survey response rates.

Based on the findings of the literature review, several design considerations for the future TTS web surveys were devised. A summary is presented in Table 14.

TABLE 14 – DESIGN CONSIDERATIONS FOR SURVEY INTRODUCTIONS AND E-MAIL INVITATIONS

Design Considerations	Reasoning
If the survey is sponsored by a government agency or an academic researcher, it should be indicated in a survey invitation or a survey's introductory statement.	• Surveys sponsored by government agencies or an academic researcher achieve higher response rates than a survey sponsored by a commercial firm.
Avoid prize and offer subject lines in e-mail invitations.	• Prize subject line and offer subject lines produce a lower response rate compared to a subject line that explicitly states the e-mail is about a survey.
Avoid personalization in invitation e-mails unless the sponsor of the survey is of high status	• Personalization tends to decrease the respondent's sense of anonymity and thus deters them from disclosing personal information in the survey • Personalization is only effective if the survey sponsor is of high statuses, such as a government agency or academic researcher. It yields higher response rates.

4.2 Progress Indicators

Some web surveys provide respondents progress bars to indicate the length of the questionnaire and the respondent's completion rate. These progress indicators can be displayed as bars illustrating the percentage of the survey completed, like the screenshot of a SurveyMonkey Research Survey shown in Figure 11, or a numerical indication of pages completed relative to the total number of pages in the survey.

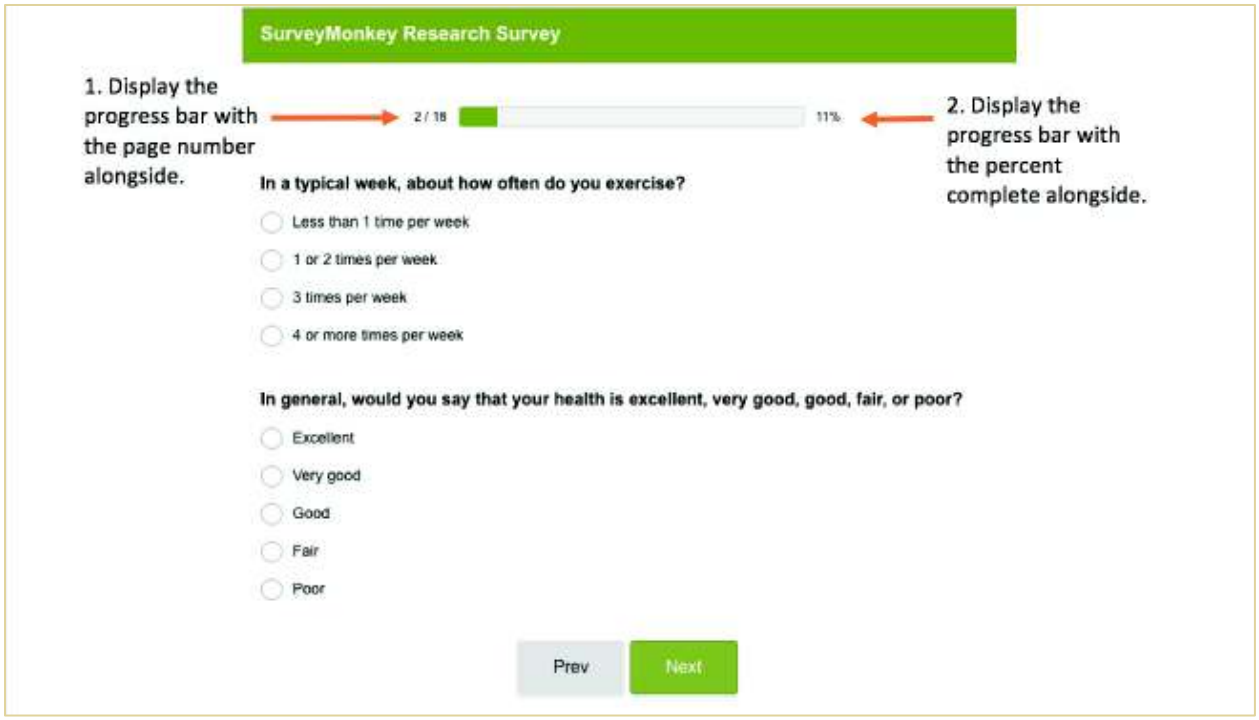

FIGURE 11 – EXAMPLE OF A PROGRESS BAR USED IN A SURVEYMONKEY WEB SURVEY (SURVEYMONKEY, 2016)

There have been conflicting results in prior research on whether progress indicators motivate or deter respondents from completing surveys. For example, Heerwegh (2004) administered surveys to 2,520 college students in Belgium and found that the progress bar decreased the drop-out rates. In addition, the results revealed that the progress bar helped decrease the percentage of missing response items for the surveys with and without progress bars had 3.2% and 4.4% missing response items respectively; however, these results were not statistically significant (Heerwegh, 2004). In a more recent study, Yan et al. (2011) compared the break-off rate for a web survey with 101 questions versus a questionnaire with 155 questions. In the shorter questionnaire, breakoff rates were 9.8% for the version with a progress bar, and 12.2% for the version without a progress bar (Yan, et al., 2011). The longer questionnaire revealed the opposite results; breakoff rates were 17.3% and 15.8% for the version of the survey with and without a progress indicator, respectively (Yan, et al., 2011). Yan et al.'s findings potentially indicate that progress indicators benefit shorter surveys more than longer surveys; however, there has been no concrete line found of when the progress indicator will yield greater or fewer breakoff rates. As the TTS is typically known to be a lengthy survey, it would follow from this latter study that the TTS should be cautious when using progress indicators. However, as mentioned before, research on presence/absence of progress indicators reports rather inconsistent findings and are very specific to the survey under study.

Crawford, Couper, and Lamias (2001) experimented with progress indicators, in addition to providing the respondents with the estimated completion times for certain tasks in a web survey. In the study, respondents were told tasks were estimated to take either 8 to 10 minutes or 20 minutes and were randomly given surveys with and without progress indicators (Crawford, et al., 2001). The results of the experiment indicated that the progress indicator tended to yield lower completion rates, particularly for respondents who were told the survey would take 20 minutes (Crawford, et al., 2001).

This finding supports Yan et al.'s (2011) argument that progress indicators benefit shorter surveys over longer surveys. Crawford, Couper, and Lamias (2001) noticed that respondents felt misled when they were told the task would last 20 minutes, but the progress indicator indicated the survey would take longer, and this resulted in higher breakoff rates.

Recall that the completion time of the TTS has varied between households, highly dependent on the number of persons in the household and their collective trips. As a result, it will be difficult to provide respondents with an estimated completion time as was done with Crawford, Couper, and Lamias' (2001) survey. Due to the inability to provide appropriate completion times, the only way to indicate progress and length of the questionnaire to the respondent is through progress indicators. Note that empirical research on the design of progress indicator (i.e. display of paging vs. percent completion vs. visual indicator on a bar) is currently lacking in the literature. Therefore, it is difficult to recommend design features to include in a progress indicator for the future TTS web survey.

4.3 Clarification and Definitions

For respondents to provide accurate and complete responses, they must possess a solid understanding of the survey questions. Often respondents misunderstand the intent of a survey question because they are unfamiliar with concepts or terms presented in the question. To avoid misunderstandings, many surveys offer definitions and clarification to key terms and concepts alongside the question instructions. However, survey respondents may think they do not need a definition for a key term especially when the concept of the question is familiar to them (Tourangeau, et al., 2013). For example, in a survey by Tourangeau et al. (2006), respondents tended to ignore the relatively technical definition of "disability" presented in the survey instructions, instead relying on their own everyday understanding of the term. Therefore, not only must the surveyor provide the needed definitions, they must also get respondents to notice and read those definitions. Conrad et al. (2010) experimented with the visibility of definitions by providing respondents with a variation of the following formats: 1) definitions always on screen, 2) rollovers which require a click/hover of a mouse to reveal the definition, 3) one-click of the term as a hyperlink to reveal the definition, 4) two clicks to reveal definition, and 5) two clicks-and-scroll to reveal definition. By using eye tracking, it was found that respondents paid more attention to definitions when they were most visible. In other words, definitions always displayed were consulted most often. Conrad et al. (2010) also argued that the length of the definitions did not affect how frequently the definition was used.

Some respondents may require additional clarification or help for a question that may not be necessary for other respondents. As an example, Conrad, Schober and Coiner (2007) found that a survey question worded as "How many people live in this house?" could be ambiguous to those in complicated living situations, such as having a child staying in a college residence. The researchers experimented with respondent-initiated clarification (i.e. requesting clarification by clicking a hyperlink) and mixed-initiative clarification (i.e. requesting clarification by clicking a hyperlink, or where clarification is offered automatically after a certain time of inactivity). The group of respondents who were not provided with any clarifications provided accurate responses for 40.9% of the questions (Conrad, et al., 2007). On the other hand, those with access to additional help improved their accuracy to 67.5% through respondent-initiated clarification and 66.4% through mixed-initiative clarification (Conrad, et al., 2007). Furthermore, the results suggested that when respondents were frequently warned they may answer incorrectly if they first did not consult the

definitions, they requested definitions more frequently (i.e. increase from 73% to 87% of definition requests) (Conrad, et al., 2007).

Based on the findings of the literature review, several design considerations for a future TTS web survey were devised. A summary is presented in Table 15.

TABLE 15 – DESIGN CONSIDERATIONS FOR SURVEY QUESTION CLARIFICATIONS AND DEFINITIONS

Design Considerations	Reasoning
Supply definitions and clarification of key terms and concepts alongside the question instructions	• Help prevent having respondents from misunderstanding the survey question
Key definitions should be clearly visible under the corresponding question's instructions	• Definitions always displayed are consulted most often than rollovers and hyperlinks
For very important definitions, respondents should be warned that they may answer incorrectly if they first did not consult the definition.	• Respondents who were frequently warned they may answer incorrectly if they first did not consult the definitions, requested definitions more frequently.

4.4 Mandatory Vs. Optional Responses, and Prompting for Responses

One benefit of web surveys over pencil and paper surveys is that respondents can be provided with instantaneous feedback/cues after an action has been performed. This is useful in indicating the completion of a question to the respondent. Also when an error occurs, the web survey can provide automatic constructive messages or cues to the respondent to rectify the error. This instantaneous feedback and ability to auto-validate responses can improve the accuracy and completeness of the survey data.

Couper, Baker, and Mechling (2011) investigated the behaviour of respondents when they were provided with mandatory and optional web survey questions, in addition to being prompted for answers that were left blank. Their study reported a 10.5% breakoff rate when questions were made mandatory, 9.4% breakoff rate when respondents were prompted for missing answers, and 8.2% breakoff rate when there were no prompts (Couper, et al., 2011). However, it should be noted that none of the differences in the breakoff rates were statistically significant (Couper, et al., 2011). Based on their findings, Couper, Baker, and Mechling (2011) suggested that the impact of mandatory responses may be small and that using prompts could be a less coercive alternative to mandatory responses. Albaum et al. (2010) conducted a similar study and found that mandatory answers increased the completeness of the data collected, and significantly affecting breakoff rates. Tourangeau et al. (2013) suggested that since mandatory responses were common in web surveys, respondents had grown accustomed to surveys that require responses; this could be the reason why studies have reported no significant correlation between mandatory responses and breakoff rates. However, as Tourangeau et al. (2013) mentioned, more research is needed on methods of optimizing

the quality of survey responses and decreasing non-responses. Surveyors should also be aware of ethical concerns that may arise for mandatory questions.

4.5 Proxy Response Biases in Trip Diaries

Currently, the TTS collects data on a household and its members by surveying a single individual of the household. This individual acts as a proxy for all other members of the household, providing travel data on their behalf. This method of proxy-reporting has always been employed by TTS, and it is a widely used method in travel surveys. Proxy-reporting comes with the benefits of faster data collection and reduced operational cost; fewer interviews are needed, and follow-up interviews are not required to contact members who were unavailable at the time of the initial interview. The Current Population Survey (CPS), jointly sponsored by U.S. Census Bureau and U.S. Bureau of Labor Statistics, report that proxy-reporting saves up to 17% of survey costs compared to interviewing all household members separately (Cobb & Krosnick, 2009).

However, the proxy-reporting method can compromise the quality and completeness of responses, as well as introduce bias in the responses. The effects of proxy-reporting on travel survey data have been analyzed and documented in various studies. This section of the report presents a literature review on the effect of proxy bias on travel studies, the demographics of proxy respondents, as well as present methods to correct for and avoid proxy bias.

4.5.1 Under Reporting of Trips by Proxy-Respondents

As mentioned earlier, the TTS has always employed the proxy interview method to collect data on entire households. The underreporting of trips in the Transportation Tomorrow Survey has also been investigated in two research papers. The first study was by Hassounah et al. (1993); they analyzed the data from the 1986 TTS, which was conducted via telephone interview. Their analysis revealed a significant difference in the overall trip rates between self-respondents, reporting on average 2.703 trips/person, and those responded by proxy, reporting 1.854 trips/person; this resulted in a difference of 0.849 trips/person (Hassounah, et al., 1993). Hassounah et al. (1993) proposed that the significant difference in trip rates was attributable to the use of proxies (i.e. differences in characteristics between the main household respondent and non-respondents of their household), and memory lapse (i.e. the incomplete knowledge the main household respondents have of the trips made by other members of their household). Not all types of trips were underreported evenly, which was clearly shown in the results of their study. They also found that the majority of underreporting trips were short, discretionary trips. Further analyses on cordon counts and TTS auto travel data were also conducted, and their results are presented in Table 16. As clearly shown, trips made during the midday off-peak period were significantly underreported than those made during the peak periods.

TABLE 16 – 1986 TTS CORDON COUNT EVIDENCE OF TRIP UNDER-REPORTING (HASSOUNAH, ET AL., 1993)

Cordon Line	Under-Reporting Rate		
	AM Peak (6:00 to 9:00)	**Off-Peak (9:00 to 15:00)**	**PM Peak (15:00 to 18:00)**
Metro-York	-3.0%	-44.8%	-8.3%
Metro-Durham	-2.6%	-45.7%	-10.2%
Metro-Peel	-2.8%	-47.5%	-9.7%
York-Durham	-3.1%	-46.8%	-10.7%
Average:	**-2.88%**	**-46.20%**	**-9.73%**

The second study was completed by Badoe and Stuart (2002), and it analyzed the 1996 TTS dataset. As with the 1986 survey, the 1996 TTS was also conducted by telephone. A significant difference was found in the overall trip rates, where self-respondents reported an average 2.818 trips/person compared to 2.235 trips/person for those responded through proxy; this resulted in a difference of 0.583 trips/person. Recall that Hassounah et al.'s (1993) study suggested that proxy respondents often failed to report approximately 1 trip (0.849 trips/person) for each of their other household members. Badoe and Stuart's reported difference was approximately 32% less than that of Hassounah et al.'s. However, their findings, as shown in Table 17, agreed with Hassounah et al.'s (1993) findings that the majority of underreporting trips were short, discretionary trips. Table 17shows proxy respondents were typically diligent in reporting their other household members' home-based work and home-based school trips, but they tended to omit a significant amount of home-based discretionary and non-home-based trips. Badoe and Stuart's (2002) also supported Hassounah et al.'s (1993) argument that trip rates are significantly underreported during the midday off-peak periods, as shown in Table 18.

TABLE 17 – 1996 TTS TRIP-RATE (TRIPS/PERSON) BY TRIP PURPOSE (BADOE & STEUART, 2002)

Respondent Status	**Home-Based Work**	**Home-Based School**	**Home-Based Discretionary**	**Non-Home-Based**
Self-respondent	1.462	1.747	1.264	0.558
Respondent by Proxy	1.492	1.783	0.766	0.247
Difference (trips/person):	**-0.03**	**-0.036**	**0.498**	**0.311**
Difference (%):	**-2%**	**-2%**	**39%**	**56%**

TABLE 18 - 1996 TTS TRIP-RATE (TRIP/PERSON) BY REPORTED START TIME OF TRIP

Respondent Status	AM Peak (6:00-9:00)	Off Peak (9:00-15:30)	PM Peak (15:30-18:30)	Off Peak (15:30-21:00)
Self-respondent	0.539	0.922	0.794	0.353
Respondent by Proxy	0.564	0.568	0.645	0.270
Difference (trips/person):	**-0.025**	**0.354**	**0.149**	**0.083**
Difference (%):	**-5%**	**38%**	**19%**	**24%**

In addition, Badoe and Stuart's (2002) study found that gender, driver license status, and household vehicle ownership were strongly correlated to the underreporting of trips. It was found that females generally took more discretionary trips (i.e. shopping, taking children to and from daycare, etc.) than males and, thus, the degree of underreported trips was greater for females (Badoe & Steuart, 2002). The results also revealed that the underreporting of trips did not vary evenly among travel modes. Across all modes, apart from public transit, there were significant discrepancies in trip rates which are shown in the results presented in Table 19. By stratifying the data by household vehicle availability (refer to Table 20), the under-reporting of trips due to proxy was more pronounced as the availability of vehicles increased for a household. Since the availability of vehicles intuitively correlates with auto trips, a high level of trip under-reporting for auto trips and an increasing underreporting of trips due to the availability of vehicles is expected. Evidently, the driver license status of a household member also correlated to the underreporting of trips, since licensed drivers are more likely to make auto trips. This is demonstrated in the results presented in Table 22.

TABLE 19 - 1996 TTS TRIP-RATE (TRIPS/PERSON) BY MODE (BADOE & STEUART, 2002)

Respondent Status	Auto-Driver	Auto-Passenger	Public Transport	Other
Self-respondent	2.070	0.310	0.310	0.128
Respondent by Proxy	1.230	0.481	0.313	0.211
Difference (trips/person):	**0.84**	**-0.171**	**-0.003**	**-0.083**
Difference (%):	**41%**	**-55%**	**-1%**	**-65%**

TABLE 20 - 1996 TTS PERSON TRIP-RATES BY HOUSEHOLD VEHICLE AVAILABILITY (BADOE & STEUART, 2002)

Respondent Status	0 – Vehicle Household	1 – Vehicle Household	2+ - Vehicle Household
Self-respondent	1.594	2.795	3.265
Respondent by Proxy	1.471	2.087	2.396
Difference (trips/person):	**0.123**	**0.708**	**0.869**
Difference (%):	**8%**	**25%**	**27%**

TABLE 21 - 1996 TTS PERSON TRIP RATES (TRIPS PER PERSON) BY DRIVER LICENSE STATUS (BADOE & STEUART, 2002)

Respondent Status	Non-Driver License Holder	Driver License Holder
Self-respondent	1.551	3.081
Respondent by Proxy	1.811	2.433
Difference (trips/person):	**-0.26**	**0.648**
Difference (%):	**-17%**	**21%**

The use of proxy-reporting in travel surveys is not a new concept, and proxy-respondents have been widely used in other household travel surveys. The Federal Highway Administration (FHWA) found that in the 1990 and 1995 National Personal Transportation Survey (NPTS), proxy-respondents reported 20% more trips and 25% more distance for themselves than for the other members of the household (Bose & Giesbrecht, 2004). An analysis of the 2001 U.S. National Household Travel Survey (NHTS) also revealed that trip rates of self-respondents were higher than the trips reported by proxies (Bose & Giesbrecht, 2004). Like Hassounah et al.'s (1993) and Badoe and Stuart's (2002) studies, Bose and Giesbrecht (2004) observed that trip rates between self and proxy responses were more significant for non-home-based trips in the 2001 NHTS. Wargelin and Kostyniuk (2014) also studied the 2001 NHTS data and found that a greater portion of proxy respondents reported no trips (i.e. 16-22% of proxy respondents reported no trips compared to 12% of self-respondents reported no trips).

In an interesting paper by Verreault and Morency (2015), they proposed that proxy bias has been decreasing over time. Verreault and Morency (2015) analyzed the 1987 to 2008 datasets from the Montreal Origin-Destination Household Survey. The Montreal Origin-Destination Household Survey is designed like the TTS, where households are contacted through telephone and one adult is interviewed per household; the only exception was the 1993 survey where interviewers were pushed to interview more than one respondent per household and were asked to ensure the collection of non-home-based trips during the day (Verreault & Morency, 2015). Evidently, 1993 data revealed a higher trip rate. Based on their analysis, they found an increasing trend in self-respondents since

1987 (Verreault & Morency, 2015). They believed that the reasoning behind this trend was due to the following demographic trends:

- Decline in household size (2.56 persons/household in 1987 to 2.38 persons/household in 2008);
- Decrease in number of children per household;
- Aging population; and
- Increase in the number of people living alone.

With an increasing number of self-respondents in household surveys, these results suggested that trip underreporting bias due to proxy have been decreasing over time and will continue to decrease until changes are observed in the triggering trends listed above.

Unfortunately, studies on tracking the effects of proxy bias over time do not currently exist for the TTS. The only studies on proxy bias in the GTA are the ones discussed in this section by Hassounah et al. (1993) and Badoe and Stuart (2002). Although Badoe and Stuart (2002) revealed a lower trip underreporting rate for the 1996 TTS compared to the 1986 TTS calculated by Hassounah et al. (1993), it cannot be concluded with great confidence that proxy bias in the GTA has been decreasing over time. Furthermore, these studies only evaluated proxy bias based on household surveys conducted by telephone. Studies on proxy bias resulting from household web surveys are missing in current literature and, therefore, this report proposes that this should be studied in future field tests; further discussion on the proposed proxy bias field test is presented in Section 6.2.

4.5.2 Demographics of Proxy Respondents

In addition to underreporting of trips, the proxy method of collecting household travel information can potentially lead to additional bias if the demographic composition of the proxy respondents is not uniform. For example, some population groups may have higher a concentration of proxy responses, and the survey data may reveal lower trip rates for these groups.

As mentioned in the previous section, Badoe and Stuart's (2002) study found gender to be strongly correlated to the underreporting of trips and, hence, underreported trips were greater for females relative to males since females took more discretionary trips. The discretionary trips mentioned by Badoe and Stuart (2002) (i.e. shopping, taking children to and from daycare, etc.) were very specific to middle-aged women. Therefore, this is a very generalized statement as it does not consider other age groups. A study by Richardson (2005) conducted a more thorough analysis of the demographics of proxy respondents based on the 2004 Coastal South-East Queensland Travel Survey (CSEQTS). Note that the 2004 CSEQTS was a self-completion survey that allowed for proxy respondents; therefore, members of the household had the option to be interviewed themselves or have another household member act as their proxy. Richardson (2005) suggested that proxy rates vary by gender as the respondents grow older. He found that young males and females (less than 14 years old) were equally likely to have their travel diary completed by a proxy (60% were completed by proxy) (Richardson, 2005). In his analysis of the older population, Richardson (2005) found that 30% of working age men (age 25 to 55) responded by proxy while only 12% of working age women responded by proxy. Overall, his study revealed that proxy respondents for males were more likely females, regardless of the male's age, whereas females were more likely to be self-respondents or have another female as their proxy, unless they were over the age of 35 in which case males were

more likely to be their proxy (Richardson, 2005). Therefore, Richardson (2005) theorized that mothers tend to fill out the surveys for their husband and children, while the husbands only fill out the survey for their wives.

Richardson (2005) also revealed that underreporting of trips varied between gender and age groups. For males, the proxy-reported trip rate was, on average, 14% lower than the self-reported trip rates (Richardson, 2005). A greater discrepancy was reported for females, where proxy-reported trip rates were, on average, 23% lower than the self-reported trip rates. Wargelin and Kostyniuk (2014) also had similar results, where proxy-reported trips by males led to a 10% decrease in trip rates, whereas proxy-reported trips by females had a 15% decrease in trip rates.

Verreault and Morency (2015) agreed with Richardson's findings, reporting that women were more often the main household respondent in the Montreal Origin-Destination Household Survey. In addition, Verreault and Morency (2015) found that persons aged 16 to 30 less often responded for their household and that there was decreasing trend over time of 16 to 30-year-old persons acting as self-respondents. They argued that the reasoning behind this trend was due to people leaving homes later in life and, thus, lowering their probability of being selected as the main household respondent (Verreault & Morency, 2015). A study by Wargelin and Kostyniuk (2014) on the 2001 National Household Travel Survey data also agreed with Verreault and Morency. Wargelin and Kostyniuk (2014) suggested that respondents possessing the following characteristics are more likely to have their survey completed by a proxy:

- Young age,
- Male,
- Student,
- Less than high school education,
- From larger household, and
- Accessible to 3 or more vehicles.

In addition, their study found that more women self-report no trips, while more men proxy-report no trips (Wargelin & Kostyniuk, 2014).

4.5.3 Correcting for Proxy Bias

The current and most widely used method for correcting travel surveys for proxy bias is the application of adjustment factors to groups of under-reported trips to match the trip rate of the self-reported trips (Hassounah, et al., 1993; Badoe & Steuart, 2002; Stopher, et al., 2003; Verreault & Morency, 2015). However, use of correction factors can become complex, especially when correcting disaggregate data sets (Verreault & Morency, 2015).

There is very little literature currently available on other methods to correct for proxy bias. Ashley et al. (2009), recommended two proactive approaches to proxy bias. The first approach is a relatively inexpensive method that involves conducting clarification interviews by phone to obtain missing or inconsistent information in the household's trip reports (Ashley, et al., 2009). The second approach is a more expensive method, and it involves using survey methods such as GPS to develop sample evidence on under-reporting, followed by the application of a correction factor to the data (Ashley,

et al., 2009). Both of these methods have been rather unpopular approaches, as they increase the response burden on the household members.

To interview all household members in a telephone survey, a convenient time needs to be negotiated between the interviewer and all household members, which is a difficult task. Therefore, the use of proxy respondents is very attractive for telephone surveys as it only requires the contact of one person from the household. In the case of web surveys, surveys can be revisited by respondents and can be completed at the respondent's time of convenience. Therefore, compared to telephone surveys, it would not be as difficult to eliminate proxy reporting and have all members of the household interviewed (apart from those who are unable such as children and elderly). The 2015 Edmonton & Region Household Travel Survey employed a mixed self and proxy reporting method for their web-travel survey where respondents had the option to independently complete their survey or have another household member complete it on their behalf; the survey is discussed more fully in Section 5.1.4. As this was a recent survey, studies on the effectiveness of this method are not yet available.

4.6 Prompted Recall Vs Announce-in-Advance Technique

The TTS, and many other travel surveyss requires respondents to report trips they have made in the last 24-hours. This is known as the recall technique, where respondents are asked to recall what happened on a prior day. The alternative to the recall technique is the announce-in-advance technique. In this latter method, the surveyor announces to the respondents ahead of time that they will have to report their trips for a specified date in the future.

Due to memory lapse, it is rather evident that the prompted recall technique can result in underreporting of trips. Various studies in the literature have investigated the degree of underreporting of trips due to the prompted recall method by comparing prompted recall survey results with GPS-based travel studies. Dumont (2009) conducted a GPS-based prompted recall survey on approximately 90 students at the University of Toronto and found that 34% reported similar trip rates to the GPS records while 53% persons recorded on average 1.78 fewer trips than in the GPS records. The remaining 13% of the participants recorded on average 1.68 more trips than the GPS records which were mostly due to participants mistakenly reporting trips for a different day (Dumont, 2009). Furthermore, Dumont (2009) found that approximately half of the participants reported an early departure time (on average 12 minutes difference), while the other half reported a late departure time (on average 16 minutes difference). A GPS household travel study at the University of Sidney (2003) and the Ohio Household Travel Survey (2002) also reported similar findings, where the number of self-reported trips was 30% less than what was captured by the GPS (Pierce, et al., 2002; Bullock, et al., 2003).

However, the comparison of the prompted-recall and announce-in-advance technique has not been an active area of research. In this literature review, it was found that many survey designers and researchers claimed that the announce-in-advance technique would produce more accurate and complete responses; however, they did not have any empirical results supporting this claim. It is believed that respondents are more likely to record their trips when they have been alerted to the fact that they will need to report their travel. Due to the lack of available research, however, the two methods should be studied in future field tests. This recommendation is detailed further in Section 6.3.

5 INTERACTIVE WEB-BASED MAPS

Currently, the TTS collects travel data from respondents primarily through Computer-Assisted Telephone Interviews (CATI). The interviewer prompts the respondent to recall details of all trips from the day before. The respondent must also relay to the interviewer the trips travelled by the remaining members of the household; this is known as proxy-reporting. The trip data collected by TTS include:

- Origin and destination locations,
- Trip Purpose,
- Arrival and departure times,
- Mode of travel,
- Details of transit trips (i.e. access and egress points and modes, routes), and
- Details of auto trips (i.e. use of Hwy 407, the number of passengers in the carpool).

Relaying such detailed information over the phone is a lengthy and cumbersome task. Apart from their home address, respondents often have difficulty providing exact addresses of places they have visited. Instead, respondents often describe locations based on closest intersections or neighboring places, which compromises the accuracy of the spatial data collected.

The TTS also has offered Computer-Assisted Web Interviews (CAWI); however, they have essentially been a direct translation of the CATI script and, thus, did not provide any additional assistance for the respondent compared to CATI.

In an attempt to relieve respondents from some burden and fatigue, several household survey administrators have taken advantage of the web's ability to integrate visual features such as maps. These integrated map tools allow respondents to search and accurately identify origin/destination locations, although they may not know the location's exact address.

This section of the report explores household surveys that have employed web-based map tools and discusses functions that could be incorporated into the map to improve the survey-taking process.

5.1 Example Interactive Map Interface Designs in Household Surveys

The use of interactive maps in household travel surveys is not a novel application. This section of the report describes how some household travel surveys use interactive maps to collect travel data. The designs of the travel web surveys discussed can serve as inspiration for a future web implementation of the TTS.

5.1.1 Utah Travel Study (2012)

Since Utah's last household travel survey in 1993, the state experienced exponential growth in population, economy and infrastructure. To analyze the impact of the state's growth and socio-economic changes on the travel behavior and patterns of its residents, a Household Travel Diary survey was conducted in 2012. Households were invited by mail to complete the survey. Participants had the option to use the online web survey or complete the survey over the phone with a trained operator. One adult household member was responsible for providing general household information, such as the number of people in the household, the number of vehicles, cumulative household income,

and demographics of each household member. Each adult household member was asked to report their trips in a 24-hour travel diary for a pre-assigned date. An adult member was responsible for completing a simplified travel diary on behalf of the minors in their household. (Resource Systems Group Inc., 2013)

The travel diary portion on the online web survey consisted of three pages: the Trip Roster page (Figure 12), the Google Map Geocoder page (Figure 13), and the Trip Details page (Figure 14) (Resource Systems Group Inc., 2013).

If the respondent had made trip(s) on their pre-assigned day, the Trip Roster page instructed them to list the locations they visited. The method of input was similar to filling in the blanks where there are prompts to the left (i.e. "I began my day at" and "Then I went to"). These prompts suggested that the list of locations must respect chronological order. Up, down, and delete buttons were also provided beside the list to allow respondents to easily rearrange the order of their trips. An example travel diary was provided on the right of the web page. The example also helped to remind the respondent to include the end location of their day as this is often forgotten. To address any further confusion that the respondent may have had regarding the Trip Roster page, a help video hyperlink was provided at the top of the page. (Resource Systems Group Inc., 2013)

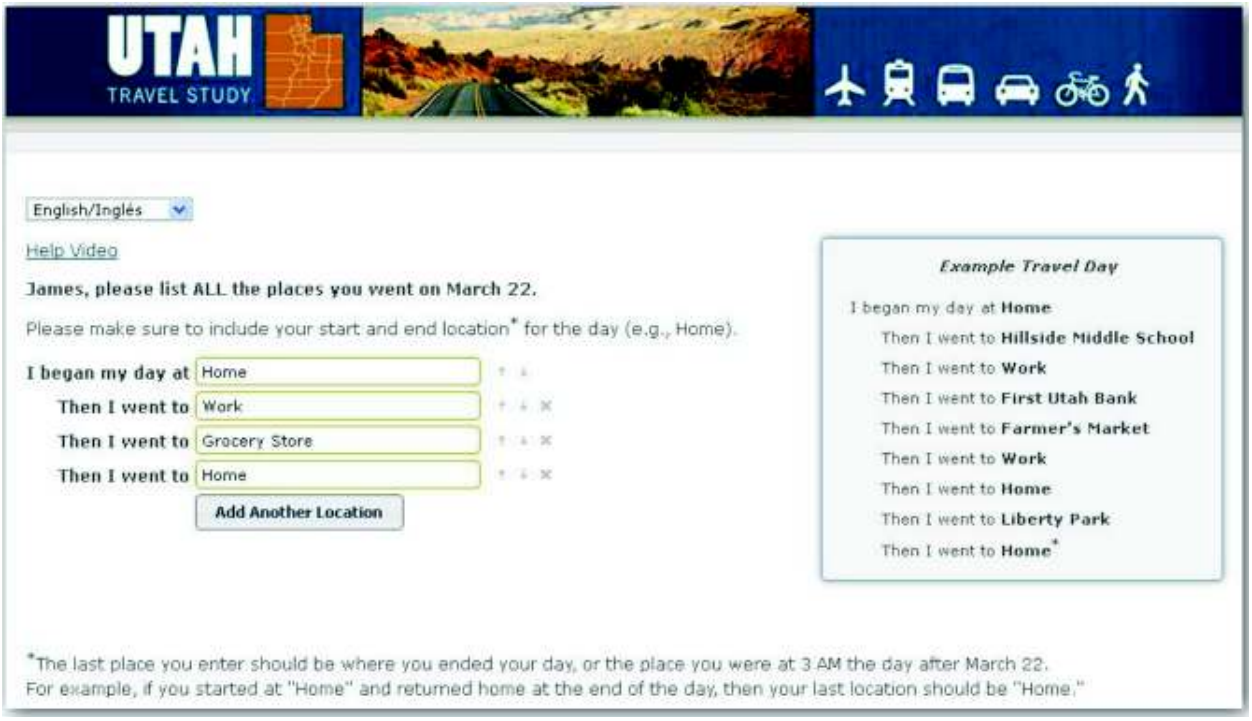

FIGURE 12 – 2012 UTAH TRAVEL SURVEY – TRIP ROSTER WEBPAGE (RESOURCE SYSTEMS GROUP INC., 2013)

After the respondent completed their trip roster, they were directed to the Google Map Geocoder page. This page instructed the respondent to locate each of the places that were listed in their trip roster. Instructions were provided to the left of the page and an interactive Google map interface was provided on the right. The respondent was first instructed to select a place from a radio button list devised from the respondent's trip roster. For the place selected, the respondent was required to

search for its location by either dropping a pin on the interactive map interface or by using a search function located below for address and businesses. To improve the accuracy of the pin drop, the pin-drop function was disabled until the respondent surpassed a certain zoom level on the map. The longitude and latitude were geocoded using Google mapping technology. The longitude and latitude were reverse geocoded into a complete address as shown to the right of the "Home" radio button in Figure 13. Once a location was identified, the address was populated beside the radio button selected, and a green check mark was displayed for confirmation. (Resource Systems Group Inc., 2013)

FIGURE 13 - 2012 UTAH TRAVEL SURVEY - GOOGLE MAP GEOCODER WEBPAGE (RESOURCE SYSTEMS GROUP INC., 2013)

Following the Google Maps Geocoder, page was the Trips Details page where respondents were asked for specific details of each trip (i.e. arrival and departure times, trip purpose, travel mode(s)…etc.). The Trip Details page walked through each of the respondent's trip in chronological order. As shown in Figure 14, a list of the respondent's trips was provided in the right-hand corner, with the trip under question highlighted. Conditionals were coded into the survey so that based on the travel mode selected, specific follow-up questions corresponding to that mode was displayed. For example, if the respondent selected automobile as their travel mode, as shown in Figure 14, the respondent was asked about the automobile use for the trip, whether the respondent was the passenger or driver, and if there were any additional costs such as tolls and parking. On the other hand, if the walking or biking travel mode was chosen, the respondent was asked different questions,

such as if they used dedicated sidewalks or bike paths. Note that trip routes data were not collected by the Utah Travel Study. (Resource Systems Group Inc., 2013)

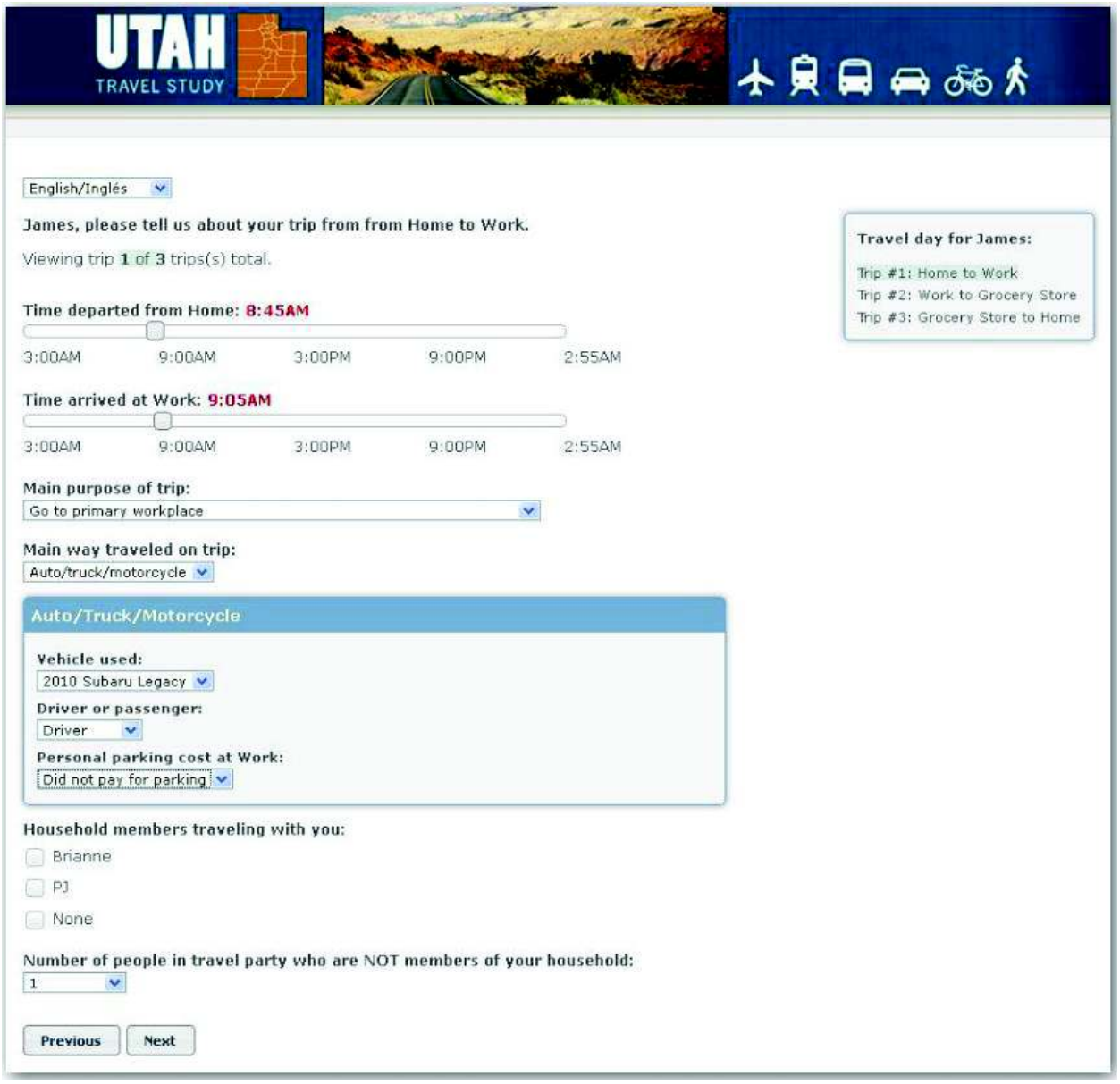

FIGURE 14 - 2012 UTAH TRAVEL SURVEY – TRIP DETAILS WEBPAGE (RESOURCE SYSTEMS GROUP INC., 2013)

5.1.2 StudentMoveTO (2015)

Currently, the TTS fails to adequately capture the travel patterns of many students. In an attempt to identify post-secondary student transportation needs, four Toronto universities (OCAD U, Ryerson University, the University of Toronto and York University) collaborated to conduct a travel survey (StudentMoveTO) targeted at their students.

Similar to the TTS, the StudentMoveTO survey requested that students provide details of their trips from the prior day. The trip diary portion of the survey began by prompting the respondent to list the places where they visited. Figure 15 is a screenshot of the survey page interface for the list of visited places. This design was referenced from the 2010 Mobility Survey of École Polytechnique de Montréal, 2011 University of Montreal's Mobility Survey and the Ministry of Transportation Quebec City Regional Survey (Bourbonnais & Morency, 2013). During the testing of the aforementioned surveys, it was found that the use of the term "places you where you went", as opposed to "trips you made", simplified the concept of trips for the respondents and caused less confusion (Bourbonnais & Morency, 2013). From the list of places, trips and trip sequences were easily devised.

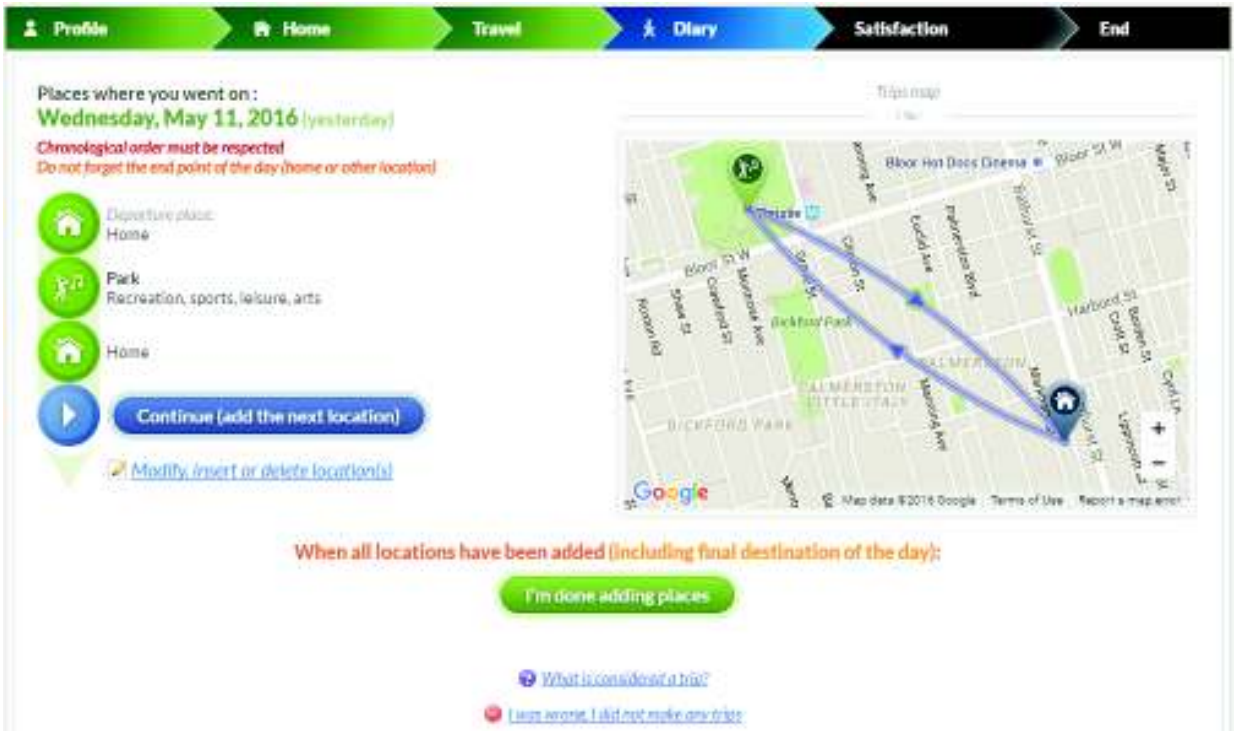

FIGURE 15 – 2015 STUDENT MOVE TO SURVEY – TRIP ROSTER WEBPAGE (STUDENT MOVE TO, 2016)

Once the respondent added a new place to their list, they were directed to an interactive map page where they were asked to identify the location of the place they added. The interactive map provided the respondent two methods to input the location: drop a pin directly onto the map, and a search function to geocode the location. As shown in Figure 16, the right side of the web page allowed respondents to geocode a location through an address search, intersection, or by selecting the location of a place the respondent had previously added.

FIGURE 16 - 2015 STUDENT MOVE TO SURVEY – INTERACTIVE MAP WEBPAGE (STUDENT MOVE TO, 2016)

Following the input of the location, the respondent was prompted to identify the main activity conducted at the location. The procedure first looped back to the list of places page to allow respondents to add additional places visited. Figure 17 shows a simplified flowchart illustrating the input sequence of the trip data for the StudentMoveTO survey. Once the respondent's list of places visited was complete, the respondent was then required to provide the arrival and departure time for each place (Figure 18), and identify the mode(s) of transport used at each trip (Figure 19). Collecting travel route information was beyond the scope of the StudentMoveTO survey.

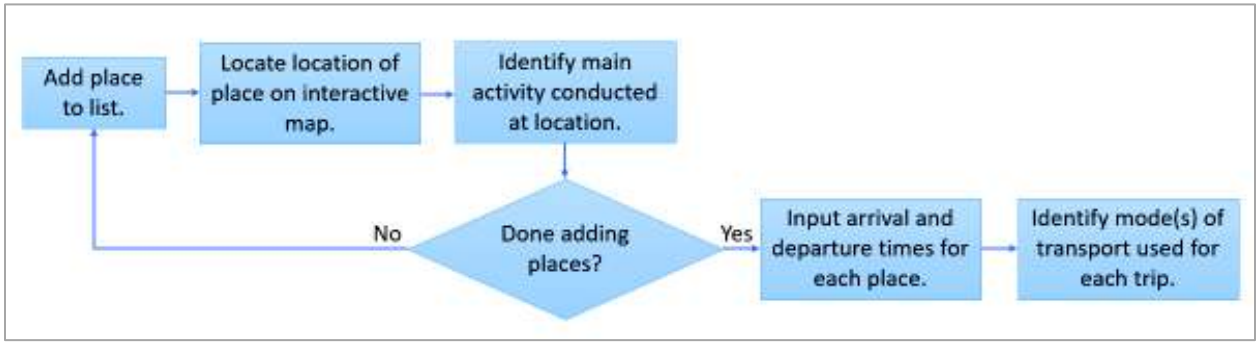

FIGURE 17 – INPUT SEQUENCE OF TRIP DATA IN THE STUDENT MOVE TO SURVEY

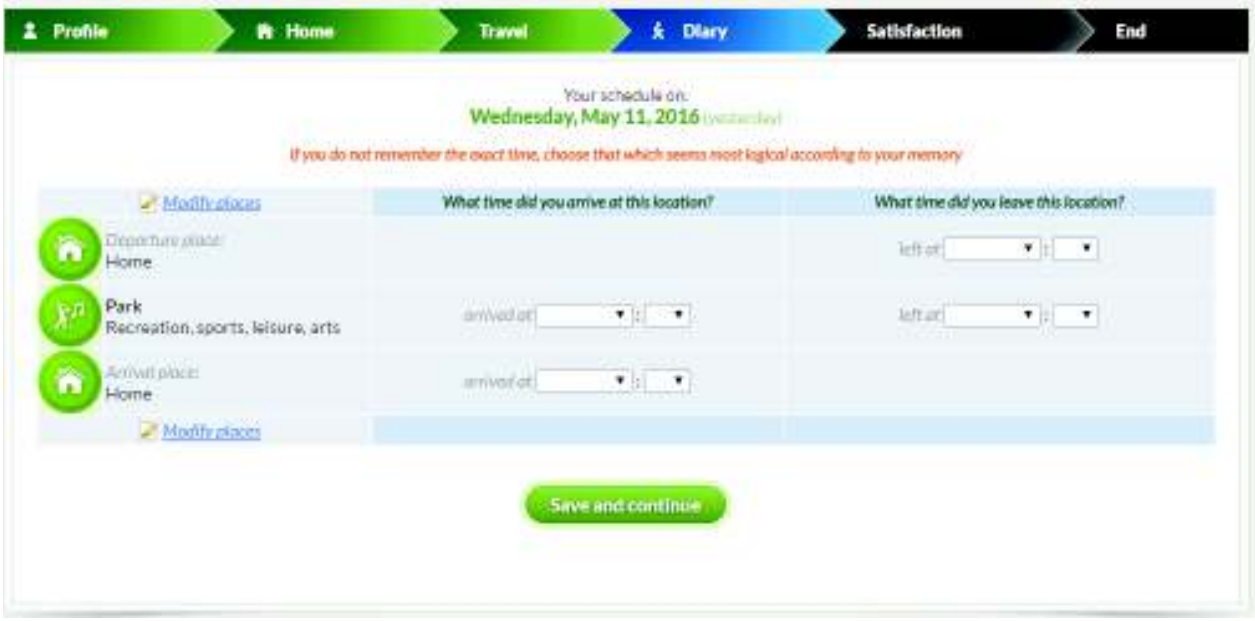

FIGURE 18 - 2015 STUDENT MOVE TO SURVEY – ARRIVAL AND DEPARTURE TIMES WEBPAGE (STUDENT MOVE TO, 2016)

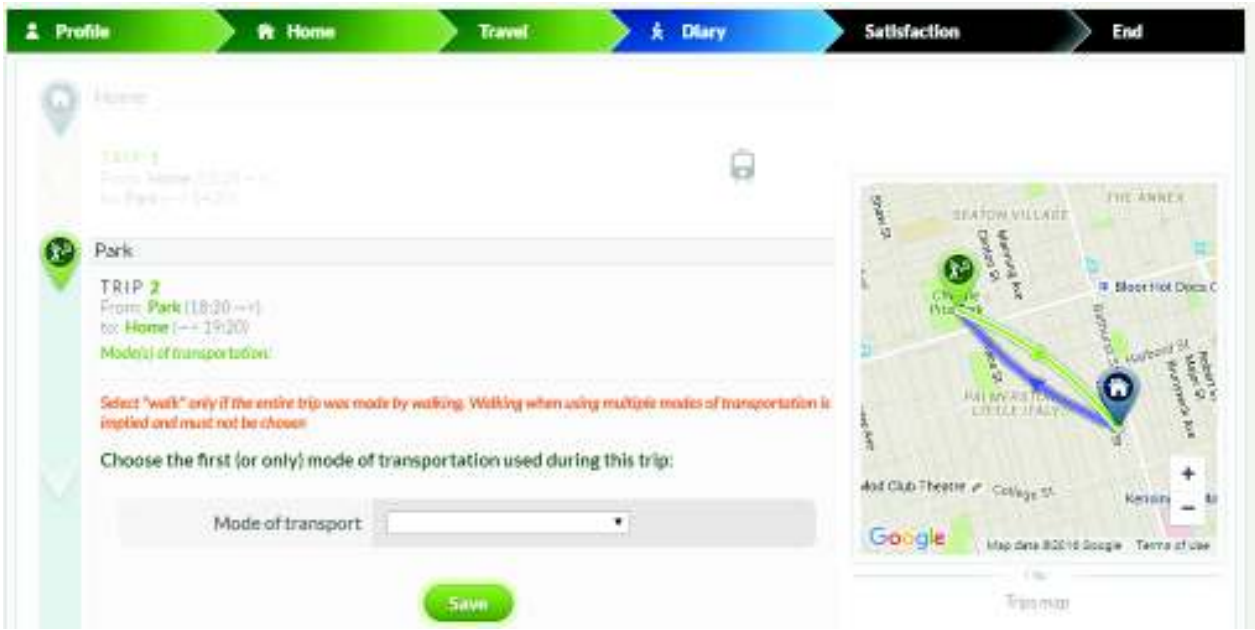

FIGURE 19 - 2015 STUDENT MOVE TO SURVEY – TRAVEL MODE WEBPAGE (STUDENT MOVE TO, 2016)

5.1.3 National Household Travel Survey (2016)

The U.S. Department of Transportation conducts the National Household Travel Survey (NHTS) every 5 to 7 years to collect travel information. Westat, a leading survey research organization that has conducted the NHTS since 2001, was also commissioned by the Federal Highway Administration to conduct the 2016 NHTS. For the web version of 2016 NHTS, Westat incorporated real-time, online

geocoding with an interactive map interface. A screenshot of the interface is presented in Figure 20. (USDOT, 2016)

A demo of the survey, or any other information, apart from the screenshot shown in Figure 20, were not available; thus, the sequence of actions and the user-interaction of the survey is uncertain.

The left section of the web page instructed the respondent to list the places they visited on their assigned travel day. Below each location listed, the respondent was prompted to provide corresponding details, such as time of arrival and departure. It appeared that once the respondent had completed the questions for one location, the box collapsed and a summary of the details was provided. The survey auto-calculated the travel time between each location and the time spent at each location. The travel time and dwell time at each location were provided to help the respondent verify the reasonability of their inputted arrival and departure times. The right half of the web page was an interactive map where respondents could visualize their trips. It was uncertain whether the map allowed the respondent to identify the location of places they listed through the pin-drop feature. Note that NTHS does not collect information on travel routes, and the routes shown on the map were just for illustrative purposes.

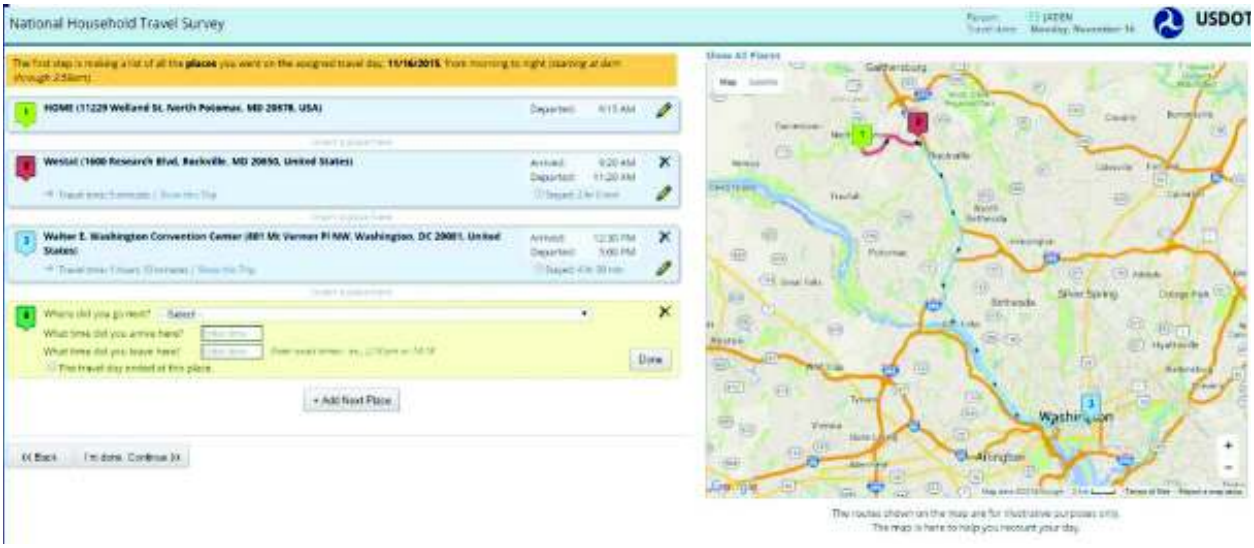

FIGURE 20 – 2016 NATIONAL HOUSEHOLD TRAVEL SURVEY SCREENSHOT (WESTAT, 2016)

5.1.4 Edmonton & Region Household Travel Survey – Making Tracks (2015)

The Edmonton and Region Household Travel Survey, also known as Making Tracks, recently contacted 253,000 households, and approximately 29,700 signed up for the survey (Edmonton, 2016). Making Tracks deemed this to be rather successful, as the region exceeded their completion targets by 14%, and had 71% of their surveys completed online (Edmonton, 2016). The Making Tracks web survey contained some novel features such as the following:

- Respondents were able save their progress on the survey, and login and logout of the survey;
- Household members were able to choose to self-report information, or to respond by proxy;
- Household were able to select a day from a prescribed list of days for which they would report their trips (announce-in-advance technique); and

- Use of an interactive web-based map to geocode origins/destinations of trips.

Similar to all other web-based travel surveys, the Making Tracks web survey requested that one adult member answer general questions about the household. In this portion of the survey, the main household member selected a travel day for which the entire household would be recording their trips. Based on the information provided, Making Tracks created a profile for each member of the household as shown in Figure 21. The screenshot of the Making Tracks household summary page (Figure 21) shows that a table of all members of the household was presented, which outlined their name, age, and gender. The table also indicated the survey completion status of each household member. Note there were two sections of the survey: individual household member's demographics, and trips. As specified on the summary page, the trips section of the survey had to be completed at the end of their specified travel day. Each household member could access the survey by using the same login the main household member, and by clicking on the hyperlinks (i.e. "Complete-Edit" and "incomplete") in the summary table.

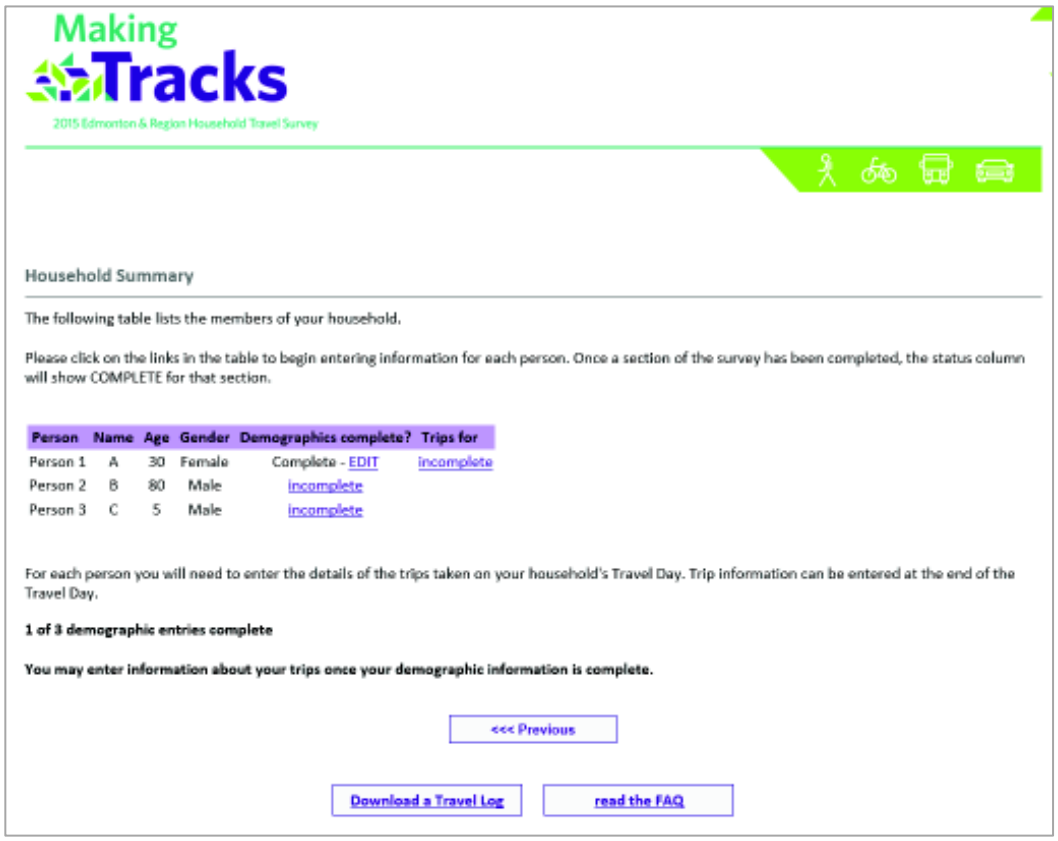

FIGURE 21 – 2015 MAKING TRACKS – HOUSEHOLD SUMMARY WEBPAGE (EDMONTON, 2016)

The trip section of the survey began by asking where the respondent was at the start of their travel day, as shown in Figure 22. The question provided radio button options of places previously entered in the survey (i.e. home, work, school…etc.), and an option to identify another location that was not listed or a "Don't know" option.

FIGURE 22 - 2015 MAKING TRACKS – TRIP WEBPAGE (EDMONTON, 2016)

Selecting the "Another location" option brought the respondent to an interactive web-based map page, as shown in Figure 23. The respondent was then able to geocode the location using the Google search bar or by dropping a pin on the map. Next, the respondent was prompted to locate the destination of their first trip, in a slight variation of the web page shown in Figure 22. The respondents were then directed to answer questions about the details of their first trip (i.e. mode of travel, arrival and departure time…etc.), followed by being prompted to input the origin and destination of their subsequent trip. The Making Tracks trip survey followed this repetitive process until the respondent had provided details on their last trip of their specified travel day. Note that the Making Tracks survey did not collect detailed route information for each trip.

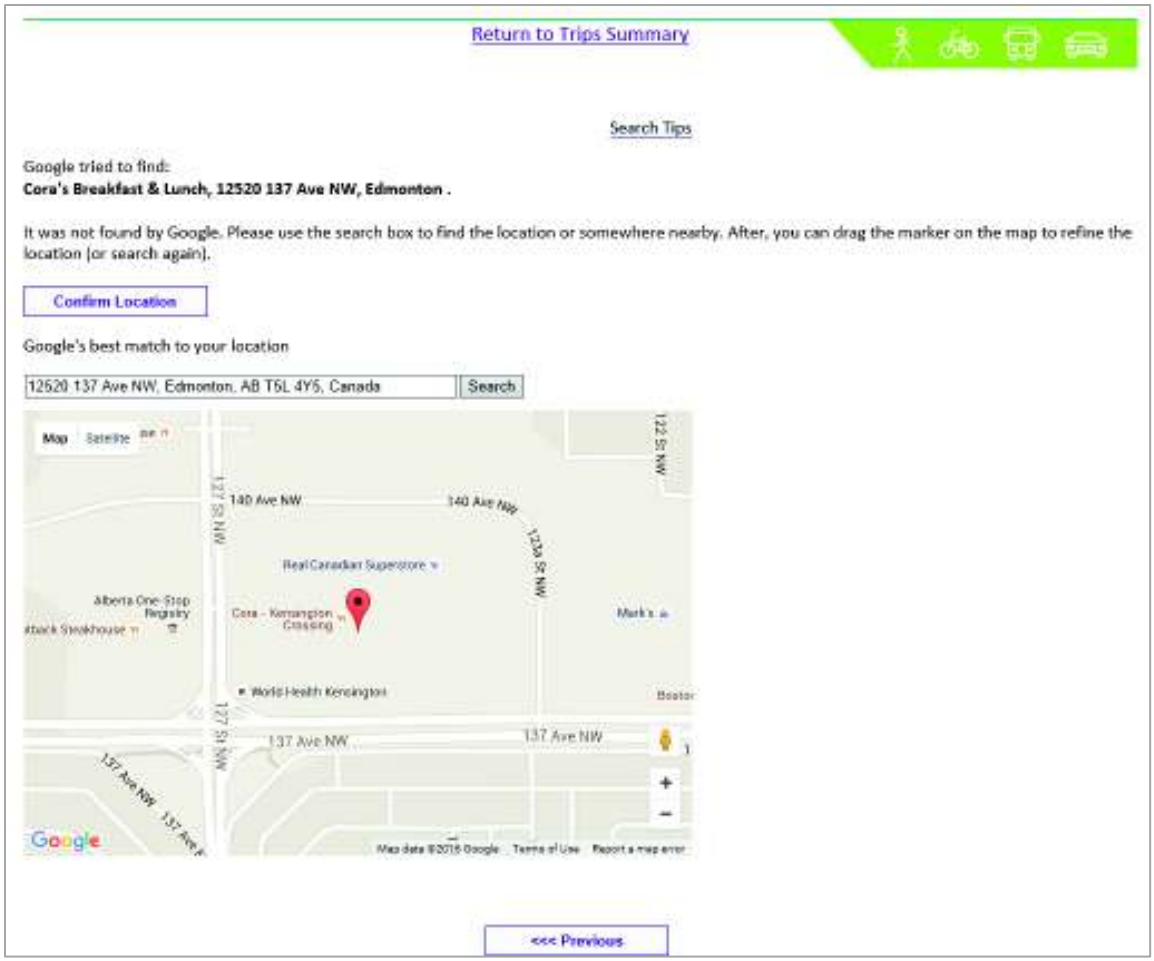

FIGURE 23 - 2015 MAKING TRACKS – INTERACTIVE MAP WEBPAGE FOR SINGLE LOCATION INPUT (EDMONTON, 2016)

5.2 Single Point Location API

One of the features of an interactive web-based map is that it allows respondents to identify a single point location; this can be done by either dropping a pin on the map or geocoding the location by typing an address, intersection or a name of a place. Geocoding is the process of translating an address to coordinates on a map. Often the search bar features auto-completion to ease the response burden. Dropping a pin on the map translates coordinates to an address, and this is referred to as reverse-geocoding. A geocoding API, such as Google Maps and Open Street Maps, allow for these functions. This feature in interactive maps comes in handy in household travel surveys, as it allows users to report locations even if they do not know the location's exact address. The pin drop method also allows for privacy; as the respondent is not required to specify an exact location, they can drop a pin in the vicinity of the privacy sensitive location.

Although interactive maps, particularly Google Maps, has been widely used by web-users, it has only been used in the more recent travel web surveys. Therefore, studies on interactive maps used in travel studies are rather scarce. However, there exists one paper by Tal et al. (2015) which analyzed web-users' behaviour towards an interactive map question in a Plug-In-Electric Vehicle (PEVs) web survey.

The survey was built on an open-source survey building platform (LimeSurvey), and it was conducted over 13 regions and states (Tal, et al., 2015). LimeSurvey currently only offers the single pin drop function; however, Tal et al. (2015) used a combination of Google APIs and the JavaScript library jQuery to implement additional spatial functionalities, such as geocoding using Google's Places API, auto-completion, and trip routing. A screenshot of the interactive map question for single location input is provided as Figure 24.

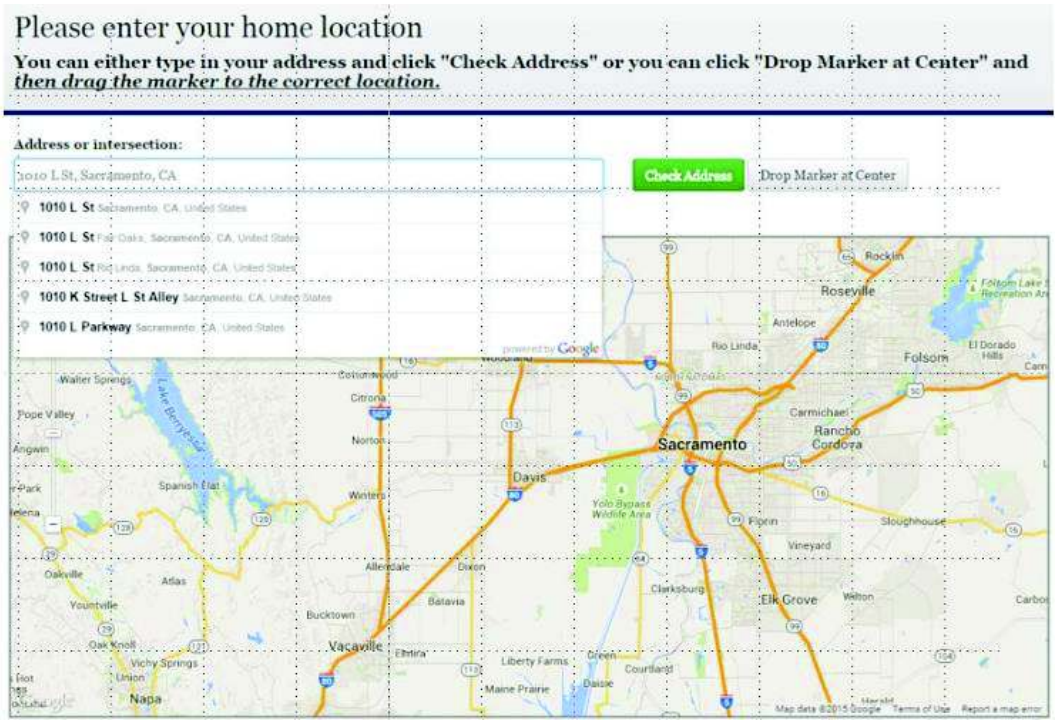

FIGURE 24 - PLUG-IN-ELECTRIC VEHICLE (PEVS) WEB SURVEY – INTERACTIVE MAP WITH SINGLE LOCATION INPUT (TAL, ET AL., 2015)

Tal et al. (2015) recorded the method used by the survey respondents for identifying their home location (i.e. pin drop versus geocoding). It was found that 81% of respondents initially geocoded their home location, and the remaining 19% of respondents initially dropped a pin on the map; the results are illustrated in Figure 25 (Tal, et al., 2015). For the respondents who initially geocoded their home location, Tal et al. (2015) also recorded whether the respondent chose an option from the autocomplete drop down menu (i.e. Google Places API) or used the geocoder directly. Tal et al. (2015) claimed that the geocoder was accurate if the respondent chose an option from the autocomplete drop down menu. Once the geocode was complete, a pin was placed on the map. To further assess the accuracy of the geocoding, Tal et al. (2015) recorded whether the respondents moved the pin or proceeded to the next question of the survey. It was found that the geocoder was very accurate, and little pin movement was required by the respondent. Similarly, for the respondents who initially dropped a pin, Tal et al. (2015) recorded the movement of the pin. The pin drop method proved to be less accurate than the geocode method; following the initial pin drop, respondents were more likely to drag the pin to another location than leave it untouched. Tal et al. (2015) also reported that for those who geocoded and dragged the pin, the median distance dragged was 1 mile on the map. For those who dropped the pin and then dragged it, the median distance dragged was greater (1.5 miles) (Tal, et al., 2015). In summary, the majority of respondents

used the geocoding feature, but if they did not, they were much more likely to subsequently drag the pin and more likely to drag it further than the respondents who geocoded.

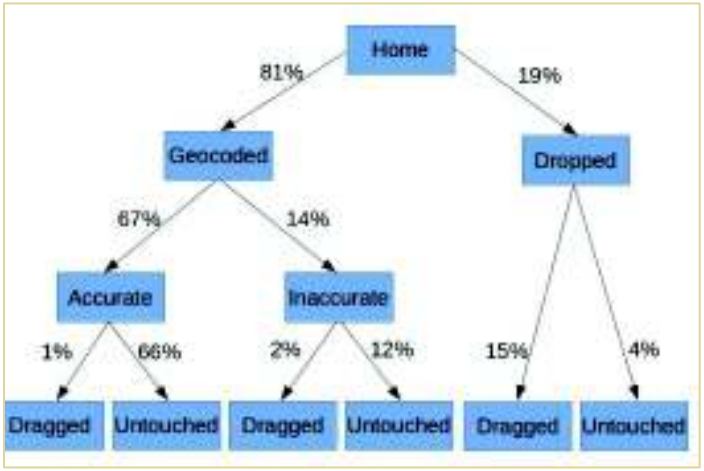

FIGURE 25 – INTERACTIVE MAP SINGLE-LOCATION INPUT BEHAVIOUR FOR IDENTIFYING HOME LOCATION (TAL, ET AL., 2015)

The survey also included an interactive map question for the respondent's work location. Figure 26 illustrates the resulting behaviour of respondents when identifying their work location on the interactive map. The results were relatively similar to the question asking for the respondent's home, except fewer people chose to geocode their work location (71%) than their home location (81%).

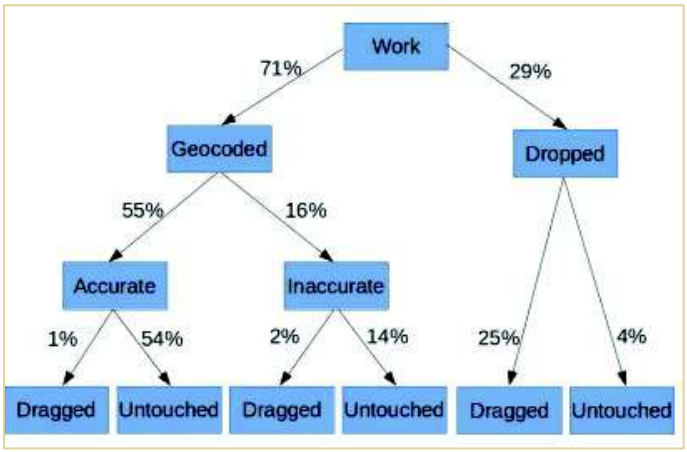

FIGURE 26 – INTERACTIVE MAP SINGLE-LOCATION INPUT BEHAVIOUR FOR IDENTIFYING HOME LOCATION (TAL, ET AL., 2015)

The data did not reveal any correlation between the method of input, the respondent choice, and the respondent's age, education, or income (Tal, et al., 2015). However, it should be noted that the survey population may have been biased, as the PEV owners tended to possess generally high income (i.e. $136,000 average for Michigan to $227,000 average for New Jersey), were better educated than the national average (nearly half of respondents hold a Master's degree or higher), and were older (median age of 52 years).

An in-house analysis of the StudentMoveTO data revealed rather different respondents' behaviour in the identifying their home location on an interactive map. A small majority (52%) of respondents

initially chose to drop a pin compared to geocoding their home location (48%). Out of the respondents who initially geocoded their home location, only 7% moved the pin, indicating that 93% of the geocodes were accurate. It could not be concluded with great confidence, however, that the other 7% of the geocodes were not accurate, as respondents may not have wanted to disclose their actual address of residence in the survey. For the respondents who initially dropped a pin, 22% of them dragged their pin on the map. A similar analysis of the StudentMoveTO data are presented in the Appendix for reference.

5.3 Trip Route API

As mentioned in the previous section, Tal et al. (2015) also recorded the trip routes using the interactive map API and was the only survey found which recorded such detail. In their survey, respondents were asked to describe a long driving trip (i.e. longer than 200 miles) that they had taken with a PEV; Figure 27 presents a screenshot of the question's interface (Tal, et al., 2015). Similar to Google Maps, the interactive map displayed the optimal route given an origin and destination on the map, and also provided alternative routes to select from. Routes displayed on the map could be manually altered by the respondent by selecting and moving waypoints along the route. Approximately 10% of the respondents selected an alternate route and 15.5% of respondents added at least one way-point to modify their route (Tal, et al., 2015). Note that the free version of Google Directions API limits the use of waypoints to eight. This did not pose a significant problem in the survey as only 0.14% of respondents used all waypoints.

FIGURE 27 - PLUG-IN-ELECTRIC VEHICLE (PEVS) WEB SURVEY – INTERACTIVE MAP WITH TRIP ROUTING (TAL, ET AL., 2015)

5.4 Design Considerations for Trip Information Collection

Since collecting trip information through web surveys is a relatively new concept, not many empirical studies on designing web survey questions for trips exist in the literature. However, there are several

papers on household travel web surveys which reflect on their experience in the design of trip questions. For example, the Regional Travel Survey for the New York Metropolitan Transportation Council (NYMTTC) collected feedback from their pilot test participants regarding their design of the trip diary question in their web survey (Chiao, et al., 2011). The participants suggested that the interface needed to look more like a paper diary (Chiao, et al., 2011). Furthermore, they wished the interface incorporated more reminder messages, and used colour and geometric shapes more efficiently to drive attention to areas where input was required (Chiao, et al., 2011).

The team who designed the University of Montreal's Mobility Survey and Ministry of Transportation Quebec Regional survey found that the following methods/features helped lower the response burden for respondents reporting trips through their web survey:

- Ask respondents to declare the set of activity locations visited on their trip day (i.e. create a complete trip roster), and then ask for the remaining details of each trip. If the trip question was designed to loop until the respondent inputs their final trip (i.e. respondent is asked about their first trip, then details of their first trip, followed by their second trip and details of the corresponding trip…etc.) respondents may experience fatigue and consciously underreport trips;
- Using the term "places where you went" instead of "trips you made" simplified the concept of trips to make sure the respondents understood what information was needed;
- Respondents should be asked if they went back home after each place they added to their trip roster in order to ensure that the returning home trip was not forgotten;
- Each time a new place was added, it can be used again as a shortcut when entering the next visited place;
- Pilot tests revealed that people tended to be more accurate when asked about what time they arrived and left a place, compared to using the terms "departure time" and "arrival time";
- A timeline helped respondents visualize their schedule of places visited; and
- Allow for multiple visits to the web survey by the respondent so they were not constrained to complete the survey in one sitting (Bourbonnais & Morency, 2013).

Although Bourbonnais and Morency's suggestions were not supported by empirical evidence or studies, intuitively they made sense, and they were supported by their expertise and experience in web survey design for travel studies. Therefore, their suggestions should be considered in the design of the future TTS web survey.

Tal et al.'s (2015) PEV survey found that their interactive map question reduced the completion rate of the survey by 4%. They realized that nearly half of the drop-outs were due to survey fatigue, unrelated to any particular question (Tal, et al., 2015). Based on their experience with the PEV survey, they provided the following recommendations for future versions of the PEV survey, which are applicable to the design of the future TTS trip question design:

- Design the interactive map question to better suit small screens (i.e. mobile devices), as the PEV survey experienced high drop rates when respondents used mobile devices;

- Conduct a browser compatibility testing to ensure quality experience for different browser users; and
- Visually distinguish between the route from the origin to the destination and the return route. Although arrows along the route were provided, the respondents did not realize the change in direction and thought the survey was directing them to the same question twice (Tal, et al., 2015).

6 RECOMMENDATIONS FOR FIELD TESTING

Based on the research presented in this report, three web-based field tests are proposed the TTS 2.0 project. The justification and recommended methods for these field tests are discussed in this concluding section of the report.

6.1 User Experience (UX) Testing

The design process of any new website or web page involves numerous user experience (UX) tests. Each web survey design is unique and, thus, its usability is difficult to predict without actual testing of the design. As mentioned before, the objective of the TTS web survey interface design is to maximize its ease of use and lower response burden, while increasing the data collected. UX testing can easily identify usability issues and provide surveyors direct feedback on how real users perceive and use the application. Furthermore, UX testing can help evaluate the user behaviour towards the interactive map question. As discussed, the use of interactive maps to collect trip information, particularly trip routes, in travel surveys is a relatively new application and very few studies exist on this subject.

Based on the research presented in this report, a list of usability measures was devised for the UX testing of the TTS web survey. Table 22 summarizes the list of usability measures and the purpose of each measure. The evaluation of these UX measures can provide an indication to the designer which elements of the TTS survey works and which areas need improvement.

TABLE 22 – UX MEASURES PROPOSED FOR TTS WEB SURVEY'S UX TESTING

UX Measures/Data	Purpose
Heat / Click Maps	• Visualize which parts of the page are frequently used. • Identify where links are expected by the user.
Mouse Movement Tracking	• Measure user's attention and predict overall experience of the user (i.e. frustration, struggling to read, uncertainty…etc.).
First & Last Device Used **First & Last Browser Used**	• Assess the usability of the survey on small screen devices by analyzing dropout rates and switching off devices • Identify if any issues experienced by the respondent is related to certain devices/browsers.
Time Stamps	• Help find most difficult parts, error prone questions and longest sections.
Selection of Route Options	• Assess how many responses relied on the route options provided.
Pin-drop vs Geocode	• Assess respondent's behavior towards the map interface. • Evaluate the accuracy of geocoding.
Way-Point Movement	• Assess the accuracy of route options provided.
Follow-up Probing Questions	• Direct feedback on the user's thoughts of the survey process

There are various methods of UX testing. Many methods involve eye-movement tracking, heat maps, sitting beside a new user using the application, and/or the use of analytic plugins. It is recommended to use the Google Analytics plugin for UX testing of the TTS web survey as it is a well-established, free program that collects the majority of the usability data listed in Table 22. The UX measures that go beyond the limits of Google Analytics can be collected directly from the survey builder platform (i.e. selection of route options, pin-drop versus geocode, way-point movements).

6.2 Proxy Bias Field Test

As discussed in Section 4.5, there has been insufficient research on the impact of proxy bias on current household travel surveys within the GGHA. The majority of the research available has either studied dated TTS data or other regions around the world. Furthermore, research available on the implications of proxy bias largely discusses proxy bias in telephone surveys. Since a key survey method proposed for future TTS is the use of web surveys, a field test evaluating the impact of proxy bias on TTS is recommended.

Given the advantages of web surveys, it is possible to survey more than one member of the household without significantly increasing the response burden. This method of mixed self and proxy reporting was used in the 2015 Edmonton and Region's Household web survey (refer to Section 5.1.4); however, the implications of this method in household web surveys has not been widely explored in research. Therefore, a field test experimenting with self-responses, proxy-responses, and mixed responses can help determine the implications these methods will have on the quantity and quality of the TTS data (i.e. response rates, dropout rates, demographics of proxy respondents…etc.). Based on the results of the proxy bias field test, it can be determined which method is most appropriate.

6.3 Comparison of Prompted Recall Vs Announce-in-Advance Technique

The comparison of the prompted-recall and announce-in-advance technique has not been an active area of research. Through the literature review presented in Section 4.6, it was found many survey designers and researchers assume that the announce-in-advance technique will produce more accurate and complete responses without the support of empirical evidence. They believe that respondents are more likely to record their trips when they have been alerted to the fact that they will need to report their travel. Therefore, a field test comparing the two techniques can evaluate the effectiveness and respondent behaviour towards each technique. Based on the results of the field test, it can be determined which technique is the most appropriate for the future TTS web survey.

REFERENCES

Albaum, G. S. & Smith, S. M., 2010. *An Introduction to Marketing Research,* New Mexico: Qualtrics Survey University.

Alsnih, R., 2006. Characteristics of Web-based surveys and applications in travel research. *Travel Survey Methods, Quality and future directions,* pp. 569-592.

Ashley, D., Richardson, T. & Young, D., 2009. *Recent Information on the Under-Reporting of Trips in Household Travel Surveys,* Melbourne, Australia: Australasian Transport Research Forum.

Badoe, D. & Steuart, G., 2002. Impact of interviewing by proxy in travel survey conducted by telephone. *Journal of Advanced Transportation.*

Baker, R. P. & Couper, M. P., 2007. The Impact of Screen Size and Background Color on Response in Web Surveys. *General Online Research Conference (GOR).*

Bell, D. S., Mangione, C. M. & Kahn, C. E., 2001. Randomized testing of alternative survey formats using anonymous volunteers on the world wide web. *Journal of the American Medical Informatics Association,* pp. 616-620.

Bergstrom, J. C. R., Erdman, C. & Lakhe, S., 2016. *Navigation Buttons in Web-Based Surveys: Respondents' Preferences Revisited in the Laboratory.* [Online]
Available at: http://surveypractice.org/index.php/SurveyPractice/article/view/303/html_51

Bernard, M., Chaparro, B., Mills, M. & Halcomb, C., 2003. Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *International Journal Human-Computer Studies,* pp. 823-835.

Bernard, M. L., Chaparro, B. S., Mills, M. M. & Halcomb, C. G., 2003. Comparing the effects of text size and format on the readibility of computer-displayed Times New Roman and Arial text. *International Journal of Human-Computer Studies,* pp. 823-835.

Boehm, L. E., 1989. *Reliability of Proxy Responses in the Current Population Survey,* s.l.: U.S. Bureau of Labour Statistics.

Bose, J. & Giesbrecht, L., 2004. Patterns of proxy usage in the 2001 National Household Travel Survey. *ASA: Survey Research Methods, Bureau of Transportation Statistics.*

Bourbonnais, P.-L. & Morency, C., 2013. Web-Based Travel Survey: A Demo. In: *Transport Survey Methods: Best Pracice for Decision Making.* s.l.:Emerald Group, pp. 207-223.

Bullock, P., Stopher, P. & Horst, F., 2003. *Conducting a GPS Survey with Time-Use Diary,* Washington, Dc: Transportation Research Board.

Chapleau, R., 2003. Measuring internal quality of a CATI travel survey. *International Conference on Transport Survey Quality and Innovation,* p. 23.

Chiao, K. A. et al., 2011. Continuous Improvement in Regional Household Travel Surveys: New York Metropolitan Transportation Council Experience. *Transportation Research Record: Journal of the Transportation Research Board,* pp. 74-82.

Childers, T. L. & Jass, J., 2002. All Dressed Up With Something to Say: Effects of Typeface Semanics Associations on Brand Perceptions and Comuser Memory. *Journal of Consumer Psychology,* pp. 93-106.

Clarke, T. & Costall, A., 2007. The emotional connotations of color: a qualitative investigation. *Color Research and Application,* pp. 406-410.

Cobb, C. & Krosnick, J. A., 2009. *Experimental Test of the Accuracy of Proxy Reports Compared to Target Report with Third-Party Validity*.. Hollywood, Florida, s.n.

Colour Blindess Awareness, 2016. *Colour Blindness*. [Online]
Available at: http://www.colourblindawareness.org/colour-blindness/

Conrad, F. G., Peytchev, A., Couper, M. P. & Tourangeau, R., 2010. Increasing Respondents' Use of Defintions in Web Surveys. *J Off Stat,* pp. 633-650.

Conrad, F. G., Schober, M. F. & Coiner, T., 2007. Bringing features of human dialogue to web surveys. *Applied Cognitive Psychology,* pp. 165-187.

Couper, M. P., Baker, R. & Mechling, J., 2011. *Placement and Design of Navigation Buttons in Web Surveys*. [Online]
Available at: http://surveypractice.org/index.php/SurveyPractice/article/view/93/html

Couper, M. P., Kennedy, C., G., C. F. & Tourangeau, R., 2011. Designing Input Fields for Non-Narrative Open-Ended Responses in Web Surveys. *Journal of Official Statistics,* pp. 65-85.

Couper, M. P. & Tourangeau, R., 2004. *Picture This! Exploring Visual Effects in Web Surveys*. [Online]
Available at: http://poq.oxfordjournals.org/content/68/2/255.extract

Couper, M. P., Tourangeau, R., Conrad, F. & Crawford, S., 2004. What they see is what they get: Response options for web surveys. *Social Science Computer Review,* pp. 111-127.

Couper, M. P., Tourangeau, R., Conrad, F. G. & Singer, E., 2006. Evaluating the Effectiveness of Visual Analog Scales. *Social Science Computer Review,* pp. 227-245.

Couper, M. P., Traugott, M. W. & Lamias, M. J., 2001. Web Survey Design and Administration. *Public Opinion Quarterly,* pp. 230-253.

Couper, M. P., Traugott, M. W. & Lamias, M. J., 2001. Web Survey Design and Administration. *Public Opinion Quaterly,* pp. 230-253.

Crawford, S., Couper, M. & Lamias, M., 2001. Web surveys: Perceptions of burden. *Social Science Computer Review,* pp. 146-162.

Dillman, D., Smyth, J. & Christian, L., 2009. *Internet, mail, and mixed-mode surveys: the tailored design method (3rd ed.)*. Hoboken: John Wiley and Sons Inc..

DMG, 2015. *TTS 2.0: An R&D Program to Develop the Next Generation Passenger Travel Survey Method for the Greater Golden Horseshoe,* Toronto: University of Toronto Transportation Research Institute.

DMG, 2016. *Transportation Tomorrow.* [Online]
Available at: http://www.transportationtomorrow.on.ca/publications.html

Dumont, J., 2009. *Trip Reporting and GPS-Based Prompted Recall: Survey Design and Preliminary Analysis of Results,* Toronto: University of Toronto.

Dyson, M. C. & Haselgrove, M., 2001. The influence of reading speed and the line length on the effectiveness of reading from a screen. *International Journal of Human-Computer Studies,* pp. 585-612.

Dyson, M. C. & Kipping, G. J., 1998. The effects of line length and method of movement on patterns of reading from screen. *Visible Language,* pp. 150-181.

Edmonton, 2016. *Making Tracks.* [Online]
Available at: https://www.edmonton.ca/transportation/traffic_reports/travel-surveys.aspx

eyequant, 2013. *108 Million Web Users Are Color Blind: How Do They See Your Website?.* [Online]
Available at: http://blog.eyequant.com/blog/2013/07/02/108-million-web-users-are-color-blind-how-do-they-see-your-website

Faulkner, C., 1998. *The essence of human-computer interaction.* London: Prentice Hall.

Fraser, T. & Banks, A., 2004. *Designer's Color Manual: The Complete Guide to Color Theory and Application.* s.l.:Chronicle Books.

Fuchs, M., 2009. The Reliability of Children's Survey Responses: The Impact of Cognitive Functioning on Respondent Behaviour. *Statistics Canada's International Symposium Series.*

Galesic, M., Tourangeau, R., Couper, M. P. & Conrad, F. G., 2009. Eye-Tracking Data: New Insights on Response Order Effects and Other Cognitive Shortcuts in Survey Responding. *Public Opinion Quarterly,* pp. 892-913.

Galitz, W. O., 2002. *The Essential Guide to User Interface Design.* [Online]
Available at: http://ps.fragnel.edu.in/~dipalis/prgdwnl/eguid.pdf

Guéguen, N. & Jacob, C., 2004. Solicitation by E-Mail and Solicitor's Status: A Field Study of Social Influence on the Web. *CyberPsychology & Behavior,* pp. 377-383.

Guner, H. & Inal, Y., 2015. The Effect of Banner Location on Banner Recognition in a Turkish Government Website: An Eye Tracking Study. *Human-Comperter Interation: Users and Contexts,* pp. 65-72.

Hassounah, M. I., Cheah, L.-S. & Steuart, G. N., 1993. *Under-Reporting of Trips in Telephone Interview Travel Surveys,* Toronto: Join Program in Transportation, University of Toronto.

Healey, B., 2007. Drop downs and scroll mice: The effect of response option format and input mechanism employed on data quality in web surveys. *Social Science Computer Review,* pp. 111-128.

Heerwegh, D., 2004. Using progress indicators in web surveys. *59th AAPOR Conference.*

Heerwegh, D. & Loosveldt, G., 2002. Online persuasion strategies. *International Workshop on Household Survey Nonresponse.*

Heerwegh, D., Vanhove, T., Matthijs, K. & Loosveldt, G., 2005. The effect of personalization on response rates and data quality in web surveys. *International Journal of Social Research Methodology,* pp. 85-99.

Joinson, A. N., Reips, U.-D., Buchanan, T. & Schofield, C. B. P., 2010. Privacy, Trust, and Self-Disclosure Online. *Human-Computer Interaction,* pp. 1-24.

Joinson, A. N., Woodley, A. & Reips, U.-D., 2007. Personalization, authentication and self dis-closure in self-administered Internet surveys. *Computers in Human Behaviour,* pp. 275-285.

Jones, R. & Pitt, N., 1999. Health surveys in the workplace: Comparison of postal, e-mail and World Wide Web methods. *Occupational Medicine,* pp. 556-558.

Kaya, N. & Epps, H. H., 2004. Color-emotion association: past experience and personal preference. *AIC 2004 Colour and Paints, Interime Meeting of the International Color Association,* pp. 31-34.

Krug, S., 2006. *Don't Make Me Think: A Common Sense Approach to Web Usability*. Berkeley: Neew Riders.

Labrecque, L. I. & Milne, G. R., 2011. *Exciting red and competent blue: the importance of color in marketing*. [Online]
Available at: http://link.springer.com/article/10.1007/s11747-010-0245-y

Labrecque, L. & Milne, G., 2011. Exciting red and competent blue: The importance of color in marketing. *Journal of the Academy of Marketing Science ,* pp. 40-45.

Leavitt, M. O. & Sneiderman, B., 2004. *Research-Based Web Design & Usability Guidelines*. [Online]
Available at: https://www.usability.gov/sites/default/files/documents/guidelines_book.pdf

Lebrasseur, D., Morin, J. P., Rodrigue, J. F. & Taylor, J., 2010. Evaluation of the innovations implemented in the 2009 Canadian Census Test. *In Proceedings of the American Statistical Association Survey Research Methods Section,* pp. 4089-4097.

Lee-Gosselin Associates Limited, 2012. *Changing Practices in Data Collection on the Movement of People,* Sainte-Pétronille: Lee-Gosselin Associates Limited.

Ling, J. & Schaik, P. v., 2006. The influence of font type and line length on visual search and information retrieval in web pages. *International Journal of Human-Computer Studies,* pp. 395-404.

Loa, P., Srikukenthiran, S., Habib, K. N. & Miller, E. J., 2015. *Current State of Web-Based Survey Methods,* Toronto: University of Toronto Transportation Research Institute.

Lynch, P. J. & Horton, S., 2001. *Web style guide: Basic design principles for creating Web sites*. New Haven: Yale University Press.

Mahnke, F. H., 1996. *Color, Environment, and Human Response: An Interdisciplinary Understanding of Color and Its Use as a Beneficial Element in the Design of the Architectural Environment*. s.l.:Wiley.

Manfreda, K. L., Batagelj, Z. & Vehovar, V., 2002. Design of Web Survey Questionnaires: Three Basic Experiments. *Journal of Computer-Mediated Communication*.

Mavletova, A. & Couper, M. P., 2014. Mobile Web Survey Design: Scrolling Versus Paging, SMS Versus E-mail Invitations. *Journal of Survey Statistics and Methodology*, pp. 498-518.

McGeeney, K. & Marlar, J., 2013. Mobile Browser Web Surveys: Testing Response Rates, Data Quality and Best Practices. *The American Association for Public Opinion Research (AAPOR) 68th Annual Conference*.

Meier, B. P. & Robinson, M. D., 2004. Why the sunny side is up: Association between affect and vertical position. *Psychological Science*, pp. 243-247.

Micheal S. McCarthy, D. L. M., 2002. Effects of typographic facturs in advertising-based persuasion: A general model and inital emperical tests. *Psycholinguistics & Marketing*, pp. 663-691.

Muehlenhaus, I., 2013. *Web Cartography: Map Design for Interactive and Mobile Devices*. Boca Raton, FL: CRC Press.

Murray, D. C. & Deabler, H. L., 1957. Colors and mood-tones. *Journal of Applied Psychology*, pp. 179-283.

Nielsen, J., 2004. *Guidelines for Visualizing Links*. [Online]
Available at: https://www.nngroup.com/articles/guidelines-for-visualizing-links/

Nielsen, J., 2007. *Banner Blindness: Old and New Findings*. [Online]
Available at: https://www.nngroup.com/articles/banner-blindness-old-and-new-findings/

Nielson, J., 2006. *F-Shaped Pattern for Reading Web Content*. [Online]
Available at: https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/

Novemsky, N., Dhar, R., Schwarz, N. & Simonson, I., 2007. Preference Fluency in Choice. *Journal of Marketing Research*, pp. 347-356.

Nygren, E., 1996. From paper to computer screen: human information-processing and user interface design. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology*, p. 188.

Odbert, H. S., Karwoski, T. F. & Eckerson, A. B., 1942. Studies in synesthetic thinking: Musical and verbal associations of color and mood. *Journal of General Psychology*, pp. 153-173.

Pearson, J. & Levine, R. A., 2003. Salutations and Response Rates to Online Surveys. *The Impact of Technology on the Survey Process*, pp. 351-362.

Peytchev, A., Couper, M. P., McCabe, S. E. & Crawford, S. D., 2006. Web survey design: Paging versus scrolling. *Public Opinion Quarterly,* p. 596–607.

Peytchev, A. & Hill, C., 2010. Experiments in mobile web survey design: Similarities to other modes and unique considerations. *Social Science Computer Review,* pp. 319-335.

Pierce, B., Casas, J. & Giaimo, G., 2002. *Estimating Trip Rate Under-Reporting: Preliminary Results from the Ohio Household Travel Survey,* Washington: Transportation Research Board.

Pope, D. & Baker, R., 2005. Experiments in color for Web-based surveys. *FedCASIC response.*

Resource Systems Group Inc., 2013. *Utah Travel Study,* Utah: Utah Department of Transportation (UDOT).

Richardson, A., 2005. *Proxy Responses in Self-Completion Travel Diary Surveys,* Washington: The Urban Transport Institute for Reliable Urban Transportation Information.

Schmidt, K. E., Liu, Y. & Sridharan, S., 2009. Webpage aesthetics, performance and usability: Design variables and their effects. *Ergonomics,* pp. 631-643.

Schwarz, N. et al., 1991. Rating scales: Numeric values may change the meaning of the scale labels. *Public Opinion Quarterly,* pp. 618-630.

Shrestha, S. & Lenz, K., 2007. *Eye Gaze Patterns while Searching vs Browsing a Website.* [Online]
Available at: http://usabilitynews.org/eye-gaze-patterns-while-searching-vs-browsing-a-website/

Sills, S. J. & Song, C., 2002. Innovations in survey research an application of web-based surveys. *Social science computer review,* pp. 22-30.

Stapleton, C. E., 2013. The Smartphone Way to Collect Survey Data. *www.surveypractice.org,* pp. Vol. 6, no 2.

Stopher, P. R., Wilmot, C. G., Stecher, C. & Alsnih, R., 2003. Standards for household travel survey - some proposed ideas. *10th Triennial Conference of the International Association for Travel Behaviour Research,* p. 24.

Student Move TO, 2016. *Student Move TO.* [Online]
Available at: http://www.studentmoveto.ca/

SurveyMonkey, 2016. *SurveyMonkey Blog.* [Online]
Available at: https://www.surveymonkey.com/blog/2016/05/17/progress-bars/

Tal, G., Favetti, M. & Nicholas, M., 2015. *Exploring the Use of a Web Based Map Tool for Travel Behavior Data Collection: Lessons from Plug-in Vehicle Owner Surveys,* Davis, CA: TRB 2016 Annual Meeting.

Toepoel, V., Das, M. & Soest, A. V., 2008. Effects of Design in Web Surveys: Comparing Trained and Fresh Respondents. *Public Opinion Quarterly,* pp. 985-1007.

Tourangeau, R. et al., 2006. Everyday Concepts and Classification Errors: Judgments of Disability and Residence. *Journal of Official Statistics,* pp. 385-418.

Tourangeau, R., Conrad, F. G. & Couper, M. P., 2013. *The Science of Web Surveys*. s.l.:Oxford Scholarship Online.

Tourangeau, R., Conrad, F. G., Couper, M. P. & Ye, C., 2014. *The Effects of Providing Examples in Survey Questions*. [Online]
Available at: http://poq.oxfordjournals.org/content/early/2014/02/17/poq.nft083

Tourangeau, R., Couper, M. P. & Conrad, F., 2004. Spacing, Position, and Order - Interpretive Heuristics for Visual Features of Survey Questions. *Public Opinion Quarterly,* pp. 368-393.

Tourangeau, R., Couper, M. P. & Galešic, M., 2005. Use of eye-tracking for studying survey response processes. *ESF Workshop*.

Tourangeau, R., Groves, R. M., Kennedy, C. & Yan, T., 2009. The Presentation of a Web Survey, Nonresponse and Measurement Error among Members of Web Panel. *Journal of Official Statistics,* pp. 299-321.

Trouteaud, A. R., 2004. How you ask counts: A test of internet-related components of response rates to a Web-based survey. *Social Science Computer Review,* pp. 385-392.

USDOT, 2016. *National Household Travel Survey,* s.l.: Westat.

UWSC, 2010. *Survey Fundamentals: A Guide to Designing and Implementing Surveys*. [Online]
Available at: https://oqi.wisc.edu/resourcelibrary/uploads/resources/Survey_Guide.pdf

Verreault, H. & Morency, C., 2015. *What about Proxy Respondent bias Over Time?,* Montreal, Quebec: CIRRELT: Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation.

Vischeck, 2015. *Vischeck simulates colorblind vision*. [Online]
Available at: http://www.vischeck.com/

W3C, 2016. *About W3C*. [Online]
Available at: https://www.w3.org/Consortium/

Walters, J., Apter, M. J. & Svebak, S., 1982. Color preference, arousal, and the theory of psychological reversal. *Motivation and Emotion,* pp. 193-215.

Wargelin, L. & Kostyniuk, L. P., 2014. *Self and Proxy Respondents in Household Travel Surveys: A Comparison,* s.l.: Research Gate.

Westat, 2016. *2015-2016 National Household Travel Survey: What's Different and New*. [Online]
Available at: http://nhts.ornl.gov/trb/2016/Workshop-Westat.pdf

Wexner, L. B., 1954. The degree to which colors (hues) are associated with mood-tones. *Journal of Applied Psychology,* pp. 432-435.

Wilkinson, S. & Payne, S., 2006. Eye tracking to identify strategies used by readers seeking informaiton from on-line texts. *Proceedings of the 13th European Conference on Cognitive Ergonomics: Trust and Control in Complex Socio-Technical Systems,* pp. 115-116.

Wright, W. D., 1988. Talking about color. *Colour Research and Application,* pp. 135-203.

Wroblewski, L., 2008. *Web form design: Filling in the blanks.* Brooklyn: Rosenfeld Media.

Yan, T., Conrad, F. G., Tourangeau, R. & Couper, M. P., 2011. Should I Stay or Should I go: The Effects of Progress Feedback, Promised Task Duration, and Length of Questionnaire on Completing Web Surveys. *International Journal of Public Opinion Research,* pp. 131-147.

7 APPENDIX

StudentMoveTO Data Analysis

Home location input method	No. of Respondents	% of Respondents
Geocode (0)	7323	48%
Pin-drop (1)	7903	52%
Total:	15226	100%

First Device Used

	Geocode		Pin-Drop		Total	
Smartphone	1210	49%	1274	51%	2484	16%
Desktop	5684	48%	6193	52%	11877	78%
Tablet	259	48%	279	52%	538	4%
Phablet	89	56%	71	44%	160	1%
Portable Media Player	14	45%	17	55%	31	0%
Unknown	67	49%	69	51%	136	1%
Total:	7323	48%	7903	52%	15226	100%

Last Device Used

	Geocode		Pin-Drop		Total	
Smartphone	1103	49%	1168	51%	2271	15%
Desktop	5811	48%	6327	52%	12138	80%
Tablet	253	49%	264	51%	517	3%
Phablet	84	56%	65	44%	149	1%
Portable Media Player	10	42%	14	58%	24	0%
Unknown	62	49%	65	51%	127	1%
Total:	7323	48%	7903	52%	15226	100%

Home pin movement	No. of Respondents	% of Respondents
Pin was moved (1)	2194	14%
Pin was not moved (0)	13032	86%
Total:	15226	100%

	Pin was moved		Pin was not moved		Total
Geocode	494	7%	6829	93%	7323
Pin-Drop	1700	22%	6203	78%	7903
Total	2194		13032		