GAMING PLATFORM VS. TRADITIONAL TEXT-ONLY STATED-PREFERENCE SURVEY OF NEIGHBOURHOOD CHOICE

Javad Mostofi Darbani

A Thesis

In

The Department

Of

Geography, Planning and Environment

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Science (Geography, Urban and Environmental Studies) at

Concordia University

Montreal, Quebec, Canada

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Javad Mostofi Darbani

Entitled: Gaming platform vs. traditional text-only Stated-Preference survey of neighbourhood choice

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Geography, Urban and Environmental Studies)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

______________________________ Chair,
Pascale Biron, PhD

______________________________ Internal Examiner,
Craig Townsend, PhD

______________________________ External Examiner,
Aaron Johnson, PhD

______________________________ Supervisor,
Zachary Patterson, PhD

Approved by: ______________________________
Chair of Department or Graduate Program Director,
Monica Mulrennan, PhD

____
Dean of Faculty,
André Roy

Date ______________________________

Abstract

Gaming platform vs. traditional text-only Stated-Preference survey of neighbourhood choice

Javad Mostofi Darbani

This research explores the influence of representational methods as they are used in stated-preference (SP) surveys of neighbourhood choice. These types of surveys have traditionally been administered in text-only format, by asking subjects which alternative they prefer based on written descriptions of neighbourhoods. It has been argued, alternately, that the visual presentation of attributes can either dominate in SP surveys, or that it can be used to help improve the realism of choice tasks, and/or increase the number of attributes that can be included in such surveys. A few studies have tested the difference between multimedia and text-only SP surveys of housing choice. While these studies have been informative, they have drawn conflicting conclusions, and have been based on small sample sizes.

The research presented here sought to take advantage of the capabilities made available by gaming engines to compare the results of SP surveys of neighbourhood choice administered either on a gaming platform, or as a traditional text-only survey. A sample of 368 (184 for each survey) respondents was used to compare neighbourhood choice model parameter estimates drawn from data administered with the two different survey methodologies. We found that while both surveys result in similar models, the model estimated using data obtained from the gaming platform show slightly better performance; providing 3D simulations appears to better focus respondent attention. At the same time, these 3D simulations have the unfortunate drawback of decreasing the generalizability of some visual parameter estimates, as they need to provide precise visual representations of their characteristics.

Keywords: Stated-Preference surveys, neighbourhood choice, attribute visualization, gaming engines

Acknowledgments

This research project would not have been possible without the support of many people. I first wish to express my gratitude to my supervisor, Dr. Zachary Patterson, who was abundantly helpful and offered invaluable assistance, support and guidance and has continually and persuasively conveyed a spirit of adventure in regard to research. Without his supervision and constant help, this dissertation would not have been possible. I am also highly thankful to Dr. Ali Rezaei for his inspiring guidance, invaluably constructive criticism and friendly advice during the project work.

Words cannot express how grateful I am to Tristan Cherry and Graeme Cherry for their support and advice upon early days of my master's degree as well as for their help on the way. Furthermore, I wish to thank my fellow graduate students at TRIP lab, Annelise Grube-Cavers, Chris Harding and Mario Perdomo for giving me valuable feedback during the evolution of my project as well as Sydney Swaine-Simon for spending many late nights programming our survey interface. Special thanks need to go to Chris Harding for editing this report; without his help this report would be unreadable.

I would like to thank my loved ones, who have supported me throughout entire process, both by keeping me harmonious and helping me put each piece together.

Contents

List of Figures

List of Tables

List of Equations

List of Appendices

1. Introduction

In analyzing and understanding people's choices with discrete choice statistical analysis, there are two principal types of data used. Revealed preference data includes observations of the choices that people have actually made. The use of revealed preference data can be challenging for a number of reasons. First, in order to undertake the analysis, not only are the characteristics of the chosen alternative required, but so are the characteristics of the alternatives not chosen. Moreover, in revealed preference data, the characteristics of the alternatives themselves can be highly correlated making it difficult to establish the independent influence of each of the characteristics. Finally, the use of revealed preference data makes it difficult to predict preferences for alternatives that do not currently exist. Stated Preference data on the other hand are obtained by asking respondents to choose between alternatives in hypothetical choice situations (choice tasks) through specially designed surveys. Since the development of these surveys requires researchers to define the alternatives, as well as their characteristics, through an experimental design, the challenges of revealed preference data can be overcome. Since they use hypothetical choice situations, they can easily incorporate alternatives that do not currently exist. As a result, these techniques have become very common in many disciplines within the social sciences and engineering (Louviere, Hensher, & Swait, 2000). This thesis concentrates on the use of these surveys in the context of neighbourhood choice.

Neighbourhood choice stated-preference surveys have traditionally been administered in text-only format by asking subjects which alternative they prefer based on written descriptions of neighbourhoods. With recent advances in computing, however, many researchers have called, and argued for, the use of visualization techniques to present certain types of information (Dijkstra, Roelen, & Timmermans, 1996; Jansen, Boumeester, Coolen, Goetgeluk, & Molin, 2009; Levine & Frank, 2006; Morrow-Jones, Irwin, & Roe, 2004; Orzechowski, Arentze, Borgers, & Timmermans, 2005; Rid & Profeta, 2011). Still, controversy remains on whether one should, and how best to use visualization techniques, and if there are advantages to providing such visual information (Arentze, Borgers, Timmermans, & DelMistro, 2003; Holbrook & Moore, 1981; Jansen et al., 2009; Orzechowski et al., 2005; Rizzi, Limonado, & Steimetz, 2012; Vriens, Loosschilder, Rosbergen, & Wittink, 1998; Wittink, Vriens, & Burhenne, 1994). The research presented here seeks to take advantage of the graphical capabilities available in gaming engines to compare the results of two stated-preference surveys of neighbourhood choice, one administered through a gaming platform, and another in text-only format.

Stated-preference surveys, also called conjoint analysis, are questionnaires designed according to statistical principles where respondents are asked to rate (or choose) between alternatives in hypothetical scenarios. Within the literature on visualization as it is used for stated-preference surveys, scholars have pointed out the following benefits. First, visualization may enhance respondents' understanding of, and decrease the ambiguity in, choice tasks (Jaeger, Duncan, & MacFie, 2001; Jansen et al., 2009). Second, it has been suggested in the literature that when some attributes are presented visually, the risk of information overload is reduced, allowing for more attributes to be included at once in a given choice tasks (Walker, Marsh, Wardman, & Niner, 2002). In line with this, Arentze et al. (2003) stated that a possible solution to decreasing burden on respondents is to add pictorial or visual information, enabling subjects to construct and maintain vivid representations of alternatives in short term memory. Third, pictorial representations of some attributes have been suspected to improve the realism of tasks, since they better mimic the actual product as experienced in the market-place. As a result of this, respondents are thought to be more likely to make the same choices in surveys as in real-world situations (Dijkstra et al., 1996; Green & Srinivasan, 1978; Vriens et al., 1998; Wittink et al., 1994).

Among the few available studies that empirically investigate the influence of representational methods, no consensus has been achieved on whether the difference in results justifies the additional time and cost associated with generating and manipulating visual materials. Rizzi et al. (2012) for example, believe that even basic images can substantially influence how attributes are perceived and valued. Therefore, they suggest that "a picture is worth a thousand words", particularly when those words describe hypothetical attributes. Conversely, Arentze et al. (2003), in a similar study, concluded that the effort involved in developing visual materials was not worthwhile, as they did not observe any influence on choice model parameter estimates resulting from visual methods of representation.

In addition to this question of time or cost benefit, some researchers have highlighted potential drawbacks from using visual information in stated-preference surveys. First, there is less control over the survey when it is presented visually. That is, some non-controlled detail in the images presented to individuals could influence their choice without it being the researcher's intent (Jansen et al., 2009). Second, attributes presented visually may then gain more importance than when presented in words (Jansen et al., 2009; Orzechowski et al., 2005; Vriens et al., 1998).

The current research is motivated by the desire to learn whether representation methods (visual and non-visual) have a significant effect on the nature and the quality of responses provided to a stated-preference survey of neighbourhood choice. To test this, two stated-preference surveys were developed; the first using the gaming engine Unity, which allowed for the creation of three-dimensional virtual environments for the gaming version of the survey, and second, a text-only version of the survey. In the gaming version, respondents were able to navigate three-dimensional simulated neighbourhoods, in which they received supplementary textual information as they explored alternative virtual neighbourhoods. By contrast, the text-only survey employed traditional means of representation so that all attributes were presented in the form of written descriptions only.

In the following section, research questions asked in the course of this project are described.

1.1. Research Questions:

This study answers questions regarding the influence of representation methods in the context of neighbourhood choice stated-preference surveys. The main research question is:

> Do the nature and quality of the results derived from a stated-preference survey of neighbourhood choice differ when responses are obtained through a gaming platform versus a traditional text-only format?

To be able to answer the main research question, more detailed questions have to be answered: Does representation method have an impact on coefficient magnitudes and their significance? Does visualization through a gaming platform enhance the understanding of respondents, and therefore decrease the amount of error variance of choice model? Does employing a gaming approach increase the predictive performance of choice models?

The research questions result from constant refinement of preliminary questions asked during this project. Similar to Arentze et al. (2003); Holbrook and Moore (1981); Jansen et al. (2009); Orzechowski et al. (2005) and Vriens et al. (1998) this study attempted to evaluate, and to compare, the gaming and text-only models in terms of coefficient magnitudes and their significance as well as the predictive performance of estimated models. Arentze et al. (2003) and Orzechowski et al. (2005), in particular, were the inspiration for investigating the presence of error variance differences across representation methods. In addition, Vriens et al. (1998) led us to ask if the nature of two models are different in terms of coefficient heterogeneity in addition to other measures.

Additionally, to validate findings of Jansen et al. (2009); Orzechowski et al. (2005) and Vriens et al. (1998) a few more questions were investigated: Compared to their written format counterparts, do attributes presented visually gain importance and then dominate? Is a gaming approach capable of providing enough level of control over details of the visual information, and therefore avoiding unsystematic influences on the choices?

The research questions will be addressed through multiple modelling approaches. First, multinomial logit models as well as mixed logit models will be estimated using the data derived from each interface to compare (1) coefficient values, (2) the significance of model coefficients, (3) the ability to account for respondent heterogeneity, and (4) the goodness-of-fit of the models. Then, a combined model with a scale parameter will be estimated to investigate the scale difference (i.e. error variance difference) between the two datasets.

Although both surveys present the same information, it is hypothesized that visual and written attributes will be interpreted differently. Specifically respondents may pay more attention to some attributes in one survey as compared to the other. This would then result in differing levels of statistical significance and coefficient values. Similarly, visualization has been argued in the stated-preference literature to be a means by which to overcome information overload in the presence of too many attributes (Arentze et al., 2003; Walker et al., 2002). Therefore, it was also hypothesized that the gaming format would yield higher levels of statistical significance for all model coefficient estimates when compared to text-only.

It was also hypothesized that the gaming format would outperform text-only in terms of predictive performance (goodness-of-fit of the choice models) and error variance (i.e. inverse of the scale parameter). This is mainly due to the fact that visualization is expected to enhance the understanding of respondents, and therefore invoke more informed (non-random) responses.

Additionally, it was hypothesized that in the gaming model, both visual and text attribute parameter estimates obtained would have sensible values. Since the graphical layout of our gaming survey provides an integrated view of visual and written information by overlaying text on graphics, respondents are less likely to fail in attending either to visual or written attributes (in contrast to findings of Jansen et al. (2009) and Orzechowski et al. (2005) where written information was ignored in some cases).

Finally, it was hoped that by using a gaming simulation, the level of control over visual attributes would be sufficient to avoid biases related to inconsequential and unsystematic information in the survey, thereby addressing concerns like those raised by Jansen et al. (2009) that these sources of unsystematic variation in the graphics would affect responses.

The remainder of this thesis starts with a review of relevant literature on the influence of representation methods in the stated-preference domain as well as previous research done within the field of housing and neighbourhood choice. Section 3 outlines the methodology employed in this research, including survey development and interface design, survey administration and the data analysis framework. The final section includes a co-authored paper submitted to the *Journal of Housing and the Built Environment*. The paper briefly reviews the previous literature and analytical framework, follows this with a discussion of modeling results, and highlights the contribution of this research. Finally, the document discusses conclusions, limitations of the current study, and possible avenues of future research.

2. Literature Review

The following section provides a review of the stated-preference literature. It is divided into four distinct parts: first, preference studies are outlined along with a brief history of stated-preference techniques; second, stated-preference-methods are described; third, the representation methods used in typical stated-preference literature are set out; the fourth part discusses stated-preference surveys within the domain of neighbourhood choice.

2.1. Preference Studies

Two types of studies, revealed and stated-preference, are typically used to investigate the factors influencing peoples' choices. Revealed preference methods use the actual decisions observed in the real-world while stated-preference methods are based on responses to specially designed surveys that put people in hypothetical choice situations (Louviere et al., 2000).

Until the 1980s, almost all policy analysis and project evaluations were done using observed choices and decisions made by individuals in real-life scenarios (i.e. revealed preference data). In these types of analyses, policy implications had to be expressed in terms of changes in current behaviour as a result in changes to the choice environment. Using observed data, however, can be challenging in some cases because of the necessity to collect information not only on alternatives that are chosen, but those that are not chosen as well. Moreover, it can be difficult to collect information or quantify data on some attributes, such as level of comfort for different modes of transportation. Finally, in cases where one wishes to study the effect of options that do not currently exist, no data is available to use for policy analysis and forecasting (Ortuzar & Willumsen, 1994).

Stated-preference methods were first developed in the field of mathematical psychology starting with the work of Luce and Tukey (1964) in 1960s (Green & Rao, 1971). In the 1970s, stated-preference methods began to be applied outside the discipline in which they had earlier developed and researchers began to experiment with transport-related choice scenarios and policy-related studies to overcome the issues associated with revealed-preference data mentioned above. These techniques permitted the study of choices and behaviour in cases where observed data could not be used. Stated-preference methods were initially discredited because there was a concern as to whether responses individuals provided in a hypothetical situation would accurately reflect the choices they would make in real-world scenarios (Ortuzar & Willumsen, 1994). However, by the 1980s, concerns on data reliability began to dissipate, after stated-preference data were compared to revealed preference data in a study by MVA Consultancy, ITS University of Leeds, and TSU University of Oxford (1987) done as part of the UK department of transport value of time project. This study resulted in a broader acceptance of stated-preference techniques, especially in transportation (Hensher, 1994).

Since, stated-preference techniques have been used in many disciplines, including for example marketing (Green & Rao, 1971; Green & Srinivasan, 1978, 1990; Wittink & Cattin, 1989), tourism and recreation (Boyle, Holmes, Teisl, & Roe, 2001; Gan & Luzar, 1993; Lin, Payson, & Wertz, 1996; Louviere & Timmermans, 1990b; Miguel, Ryan, & McIntosh, 2000), agricultural economics (Gillespie, Taylor, Schupp, & Wirth, 1998; Harrison, Stringer, & Prinyawiwatkul, 2002; Holland & Wessells, 1998; Prentice & Benell, 1992), and transportation studies (Hensher, 1994; Louviere et al., 2000) .

Moreover, from a modelling point of view. the explanatory variables in real world scenarios do not always have the range of values or variability desired in order to enable proper estimation of their effect on choice. Therefore, it may be challenging to develop behavioural models that can be used to predict responses to changes in the status quo due to a new policy. It is thus often necessary to design stated-preference surveys to provide insight into the likely market response to new policies. these surveys have the advantage that they can be designed to avoid, or at least mitigate, problems associated with revealed preference data (Ortuzar & Willumsen, 1994). Revealed and stated-preference techniques are, however, complimentary and there has been growing interest in combining the two data sources in transportation, marketing and environmental studies during the past decade (Louviere et al., 2000).

Last but not least, the stated-preference method is a tool that can be used to inform policy makers about what the public prefers and values most in community development projects and transportation studies. New urban projects mostly are accompanied by estimate of cost over time and revenue forecasts to see if the benefits and outcomes are worth the huge investment required to bring a project to fruition (Ortuzar & Willumsen, 1994). Moreover, the technique has been found to be an appropriate tool for assessing monetary values for improvements to the environment and to predict responses of a target group to development policies, such as introduction of new urban infrastructure or a new housing alternative (Louviere et al., 2000; Ortuzar & Rodríguez, 2002; Rid & Profeta, 2011).

2.2. Stated-Preference methods

"Stated-Preference" techniques can actually refer to a number of different survey techniques used to understand people's preferences. The three most common stated-preference methods techniques are contingent valuation (CV), conjoint analysis (CA) and stated choice (SC) techniques. Contingent valuation deals with estimating respondent willingness-to-pay for a particular policy or product option, while conjoint analysis and stated choice allow the researcher to study preferences and willingness-to-pay for the entire policy or product option, as well as for its individual characteristics (Ortuzar & Willumsen, 1994) .

Stated-choice and conjoint analysis are similar as both put the respondent in a hypothetical situation. However, in conjoint analysis respondents are asked to rank a series of alternatives at the same time. In contrast, in stated-choice surveys, respondents are asked to choose their preferred alternative from a subset of hypothetical alternatives chosen carefully by the researcher. Moreover, in a stated choice survey (sometimes referred to as a Choice-based Conjoint Experiment, or Discrete Choice Experiment) respondents are asked to make a number of such choices. In the field of transportation, the stated-preference label has referred to both conjoint analysis and stated choice, with no formal distinction (Louviere, Flynn, & Carson, 2010; Ortuzar & Willumsen, 1994; Ramírez-Hurtado, 2010). In the rest of this thesis, the term "Stated-Preference" survey is used in the sense of a stated choice survey.

As such, stated-preference surveys, are a data collection approach where respondents are asked to choose between alternatives in specially designed hypothetical scenarios where alternatives are characterized by attributes of different levels. These scenarios are referred to as Choice Tasks (Louviere et al., 2000). Task complexity, in terms of the number of attributes used to define the alternatives and their associated range, the number of choice situations presented to each respondent, and the representation method used to convey information to respondents are all important factors considered when developing a stated-preference survey. Lack of attention to any of these aspects may have important consequences, and may affect the response quality (Levine & Frank, 2006).

Therefore, the design of a stated-preference survey demands careful attention to many details, and may be subject to several sources of error. Firstly, the number of attributes should be relatively small to keep the survey design manageable (Hunt, McMillan, & Abraham, 1996; Walker et al., 2002). On the other hand, omission of an important attribute could lead to inaccurate estimation of the relative importance of the attributes included (specification bias). Therefore, the decision about which attributes to include is very important. Secondly, ensuring an understanding of attributes among respondents is another important aspect of stated-preference survey design; the manner in which attributes are presented could potentially lead to misunderstandings. As a result of this, the validity of responses would then be called into question. This is referred to as instrument bias (Kim, Pagliara, & Preston, 2005).

2.3. Representation methods in Stated-Preference surveys

The question of how best to provide information in stated-preference surveys has received a fair bit of attention in the literature. Few studies, however, have tried to empirically evaluate the influence of different representational methods (Rizzi et al., 2012; Wittink et al., 1994). In the absence of empirical evidence about the effect of visualization on results of stated-preference surveys, text-only representation methods have been considered adequate by most of the researchers, and surprisingly, little research has been done in this respect (ibid). In this section available literature concerning the effect of different representational techniques, both within the broad stated-preference survey design domain, as well as that which regards housing and neighbourhood choice specifically (Arentze et al., 2003; Dijkstra et al., 1996; Jansen et al., 2009; Orzechowski et al., 2005; Rizzi et al., 2012; Vriens et al., 1998; Wittink et al., 1994) are discussed.

A number of studies have identified benefits to visualizing attributes. First, visualization may enhance respondent understanding and decrease the ambiguity of choice tasks, which in turn can lead to more heterogeneity in responses. Vriens et al. (1998) for example found higher response heterogeneity when images were used in stated-preference surveys. That is, they were able to have more segments for one of the visual attributes in the latent class model they developed, concluding that this may be the result of a more informed evaluation by the respondents in the visual representation sample.

Second, it has been suggested that more attributes can be included in choice tasks while reducing the risk of information overload, when some attributes are presented visually (Wittink et al., 1994). In line with this, Holbrook and Moore (1981) found more significant variables in their stated-preference surveys thanks to the inclusion of simple line drawings in their surveys of sweater choice (see Appendix 3 for an example of the choice task). This, however, is not a universal finding; Vriens et al. (1998) for example found no difference in the number of significant attributes in a similar study related to car stereo design (see Appendix 2 for an example of choice task from Vriens' et al. study).

Third, pictorial representations of some attributes have been expected to improve the realism of tasks, since they can better mimic the actual evaluation of a product in the market-place, and therefore that respondents would be more likely to make the same choices in the surveys as in real-world situations (Dijkstra et al., 1996; Green & Srinivasan, 1978; Vriens et al., 1998; Wittink et al., 1994).

One of the most recent studies concerned with visualization is done by Rizzi et al. (2012) investigating the impact of traffic images on value of travel time savings within a route-choice experiment. A total of 481 respondents were randomly assigned to one of two surveys: the first made use of a text-only approach, while the other provided images of congestion and free-flow travel conditions along with written information. See Appendix 1 for an example of a choice scenario with traffic images. The text-only survey elicited the same value of travel-time-savings for traveling in a congested and free-flow traffic conditions. Conversely, the survey with images revealed that respondents perceiving a 30% additional value for their time for traveling in congested conditions. That a higher value should be placed on time spent in congested traffic conditions, according to the authors, is commonly accepted in recent mode choice literature, and the authors concluded that even basic images seem to substantially influence how traffic conditions are perceived and, therefore, how these travel times are valued. The work of Rizzi et al. (2012) provides preliminary evidence that the cost and effort of including images in stated-preference surveys may be worthwhile, and their piece concludes by suggesting that more research is necessary to understand the effect of incorporating sophisticated traffic imaging, such as real time images into stated-preference surveys.

In contrast to the study done by Rizzi et al. (2012), a similar study by Arentze et al. (2003) casts doubt on whether representational methods that use images significantly influence the result of stated-preference surveys. Arentze et al.'s study investigated the effect of representation methods, as well as task complexity in terms of the number of attributes, alternatives, and choice sets as well as literacy of the respondents in a stated-preference survey concerned with transport mode choice. The authors found no statistically significant difference in error variance of the model (scale parameter difference), or in the attribute weights were found across the two representational methods (pictorial versus text-only description). However, they did find substantial differences in error variance when the number of attributes was increased from three to five. Based on these findings, the authors suggest it is not worth the extra effort and cost of developing visual materials for stated-preference surveys, as no difference in results was observed between the different representation methods. They further suggest that it is more worthwhile to focus attention on other aspects of the design of stated-preference surveys, such as the number of attributes and their associated range as well as choice situations presented to each respondent.

With respect to the use of visual information in housing and neighbourhood stated-preference surveys, a few studies were identified. Jansen et al. (2009) conducted two studies to examine the impact of including images along with text in stated-preference surveys of housing choice. In the first study, three different versions of their survey were used: one text-only version, another containing text with black and white drawings of alternative houses, and a third with color photos with text. The surveys had 5 attributes and were completed by the same 28 respondents. In their second study, they compared a text survey with photos available on request by the respondents and another with text and photo montages. In this second study, both surveys consisted of 13 attributes, 7 attributes related to dwelling characteristics and six for characteristics of the dwelling environment. 48 respondents filled out the text survey and 59 completed the survey with text and photomontages. Appendix 4 and Appendix 5 show examples of the survey instruments. Jansen et al. (2009) found differences in estimated parameters between the stated-preference surveys that they hypothesized could be the result of unsystematic variation and non-controlled details in images presented in the surveys with images. They also found that visually presented attributes were assigned a greater importance when compared to the same attributes presented in text form.

Orzechowski et al. (2005) conducted a study related to housing choice that compared a multimedia representation methodology (Virtual Reality) against a traditional text-only survey. To the best of my knowledge, this study is the only application of virtual reality in the use of stated preference surveys applied to the field of housing. The multimedia version of the survey allowed respondents to change their perspective within a simulated environment depicting architectural plans for houses with different layouts. Four attributes were used in these multimedia surveys, with price being the only attribute presented in text format.

Appendix 6 shows an example of the two survey instruments. No statistically significant difference was found in terms of the internal and external validity of two representation methods. The coefficient estimates for the price attribute, however, were not reliable in the multimedia survey. The authors suggest that this was because the visual attributes dominated when contrasted with attributes presented using text. Finally, the authors expected less variance of the error term in models developed using data collected with the multimedia representation, since the graphical presentation of attribute levels was assumed to make the attribute interpretation easier. This expected result was not borne out in the results however; while the visual survey models did have a lower error variance, the difference in error variance was not statistically significant. This result may have been due to the fact that there were relatively few respondents to the survey, having only 35 respondents for the text-only survey and 29 for the multimedia survey.

To summarize, few studies have attempted to compare the dissimilarities obtained when using different representation methods with choice tasks. This is striking, especially in comparison to the considerable literature available on other dimensions of survey design. The review of literature shows there is no consensus in the literature on whether inclusion of images in stated-preference surveys is worthwhile or not. Authors have used actual prototypes of products, realistic pictorial presentations and 3D images to explore the question of whether visual modes of representation are superior to traditional text-only method. Despite the benefits pointed out in the literature for visual representations, past studies seem to have had trouble ensuring systematic control of visual attributes (i.e. with Jansen because of the use of pictures, renderings and collages of pictures), had small numbers of attributes defining the choice subjects to effectively understand the effect of visualization (i.e. as in Arentze's study), or failed to properly present attributes in an interface where respondents adequately analyzed both visual and written information (as was the case in Jansen et al. (2009); Orzechowski et al. (2005)). As can be read by the issues highlighted above, it is evident that there is still room for further research in this respect.

The next section reports on stated-preference studies within the housing and neighbourhood choice domain. These stated-preference surveys were used as a guide for selection of written and visual attributes for our research project.

2.4. Stated-Preference and Residential neighbourhood choice

Stated-Preference (SP) surveys in the domain of housing and neighbourhood choice have typically been presented in the traditional text-only format (Cooper, Ryley, & Smyth, 2001; Hunt, 2001; Hunt et al., 1996; Kim, 2006; Kim et al., 2005; Louviere & Timmermans, 1990a; Ortuzar, Martinez, & Varela, 2000; Senior, Webster, & Blank, 2006; Walker et al., 2002; Wang & Li, 2004). There have been, however, a few studies that have applied multimedia approaches (i.e. a combination of written and visual representations) to stated-preference surveys (Jansen et al., 2009; Levine & Frank, 2006; Morrow-Jones et al., 2004; Orzechowski et al., 2005). In this section, previous studies in the domain of housing and neighbourhood choice are reviewed.

With respect to the traditional approach to stated-preference surveys in housing and neighbourhood choice, Hunt (2001) is an often cited example, and is perhaps the most comprehensive in a Canadian context. The author's work examines the impacts of transportation-related factors on the attractiveness of residential locations in Edmonton, Canada. As Molin and Timmermans (2003) state, this study could be considered a further development of his earlier work, where the author put together a comprehensive list of attributes considered influential in housing choice behaviour by reviewing past revealed-preference and stated-preference studies.

Hunt's earlier work, Hunt et al. (1996), limited the number of attributes to only five characteristics assumed to be of highest importance when selecting among residences; cost per month, number of bedrooms, minutes of in-vehicle travel time to work, minutes of in-vehicle travel time to a shopping center, proximity to a light rail transit station. Each alternative was presented on a separate card and respondents were asked to rank four randomly selected cards from best to worst. The survey was conducted in Calgary, Alberta and a total of 390 surveys were done. In Hunt (2001) however, nineteen variables were combined into each housing alternative. The attributes were organized into eight related groups, and four of the attributes were held constant across alternatives to decrease the task complexity; see Appendix 7 for an example of choice task from Hunt (2001). A random experimental design was used except for when the combination of attributes could not possibly or reasonably occur. In order to establish a consistent understanding of the attributes, supplemental materials were provided in the form of pieces of paper with bulleted points and in some cases photographs. Interviews were scheduled via telephone, and surveys were administered face-to-face due to survey complexity. The total number of respondents consisted of 1,277 interviews. A logit choice model was used for analysis of the survey responses.

Gayda (1998) conducted a stated-preference survey of residential location choice in Brussels, Belgium with the purpose of examining and quantifying the trade-offs between important attributes of residential location. As it was administered by mail, the survey was designed to be simple, to guarantee a reasonable response rate and high quality for collected data. As such, it consisted of only two choice scenarios where each scenario had two alternatives. The survey involved only three attributes, namely price of housing, type of neighbourhood in terms of level of traffic, and duration of home-to-work journey. The survey was designed in two stages. First, a survey was mailed to respondents to collect data on household characteristics. Then, this information was used to design customized stated-preference questionnaires for each household; the dwelling prices, in particular, were customized to account for the maximum budget each household could allocate to housing. 429 valid responses were received and a multinomial logit model was used for estimation.

Kim et al. (2005) implemented a stated-preference study to explore the extent to which transport and neighbourhood amenities affect the decision to move and the choice of residence, including dwelling and location. Considering the limitation of humans' cognitive ability, two stated-preference surveys were set up in this study, each having five attributes, but where both had housing price as a common attribute. The first survey was intended to understand the trade-offs between house prices and transport attributes (travel time to work, travel cost to work location), while the second survey was designed to capture the trade-offs between house prices and location amenities (population density, travel cost to shop, school quality). The surveys were conducted in Oxfordshire, UK. Each of the two stated-preference surveys was made up of 16 choice tasks with three alternatives. The final sample consisted of 96 usable self-completed questionnaires, resulting in 1,536 observations for each. The two datasets were combined to give 3,072 independent observations used in estimating a nested multinomial logit model.

Kim (2006) conducted another stated-preference study investigating the marginal value of housing attributes in residential location choice in the Ulsan metropolitan area in Korea. This study analyzed data collected through two separate stated-preference survey with five attributes. Citing Walker et al. (2002), the authors argued that the number of attributes should be limited to no more than five due to the limitation of human cognitive abilities. The two surveys had two common attributes (house price, travel time to work) and three different attributes (travel cost to work, noise and access to park within the first survey, and travel time to shop, dwelling size and access to school within the second survey). Each of the two surveys consisted of 18 choice tasks completed by 188 owner-occupied households through face-to-face interviews. A multinomial logit approach was used to estimate two separate models for each data set and one for the combined data set. The repeated nature of the responses was not taken into account, however, a relatively sophisticated combined estimation approach of data was used in this study. The authors estimated two separate and one combined model with data collected through the two surveys. As neither the magnitude of the two common coefficients nor the error variance of all the three models (i.e. scale parameter) were statistically significantly different, they concluded that all three models were comparable.

With respect to the use of visual approaches in stated-preference surveys of neighbourhood preference, Levine and Frank (2006) sought to elicit individuals' views on attributes of neighbourhoods. The purpose of their research was to measure the desire for various neighbourhood types, from preferences for low-density, auto-oriented environments to desires for compact, walkable and transit-oriented neighbourhoods in metropolitan Atlanta, USA. 1,455 randomly selected households were asked to rank two competing statements describing particular neighbourhood alternatives. See Appendix 8 for an example of the preference tasks. This study was not concerned with the representational issues of stated-preference surveys, but was of interest to our study because it emphasized the necessity to include sketches of neighbourhood designs in order to properly convey information about either density or land use separation. The authors argued that past residential preference surveys suffered from a lack of visual representation of urban design concepts.

The review of the literature shows, on one hand, that the housing and neighbourhood attributes that need to be included to make these studies useful and realistic are numerous. As a result of this, many past researchers have attempted to incorporate as many variables as possible by employing different design strategies. For example, grouping variables, as in Hunt 2001 or developing two separate but related stated-preference surveys as in Kim (2006) and Kim et al. (2005). At the same time, some authors (Dijkstra et al., 1996; Jansen et al., 2009; Levine & Frank, 2006; Orzechowski et al., 2005) have been interested in the visual representation of attributes in these surveys.

Visual materials have also been used in the context of stated-preference surveys, either as supplementary information, as seen in Hunt (2001), or in the survey design itself, as done in Jansen et al. (2009); Levine and Frank (2006); Orzechowski et al. (2005). All studies reported here were used to identify potential attributes to be included in our stated-preference surveys but also highlight that, to date, text-only representation has predominantly been used in the literature. What the work presented here achieves is to fill the gap with respect to developing an interface that not only enhances the understanding of respondents by providing visual information, but also one that allows respondents to adequately attend to both visual and text attributes.

3. Methodology

This section discusses the process of setting up current stated-preference study. First, survey interface development and design are discussed. Then, survey administration and sampling methodology are described. Finally, the data analysis framework is covered in detail.

3.1 Survey Interface Development and Design

Survey interface development is an important part of any stated-preference study, and lack of attention to any of the steps described below may have important consequences and affect the quality of responses (Levine & Frank, 2006). Although designing a stated-preference survey is an iterative process it still needs to follow a number of steps (Louviere et al., 2000):

1- define of study objective;
2- conduct supporting qualitative study for selection of attributes;
3- develop the data collection instrument;
4- define sample characteristics;
5- collect data;
6- analyze collected data and estimate model;
7- conduct policy analysis;

3.1.1. Study objective

The objective of this study is to compare the statistical results obtained from two stated-preference surveys with different representational methodologies. To answer the research questions, two stated-preference surveys of neighbourhood choice, one using a gaming approach and another with text-only means of representation, are developed and administered. Since this study is only concerned with the influence of representational methods on statistical model estimation, policy analysis was deemed beyond the scope of this project.

3.1.2. Selection of the Attributes for the Stated-Preference Surveys

The first step in designing any stated-preference survey is determining the most salient attributes influencing the choice behaviour under study. Attributes might be selected through qualitative research, based on experts' experience, through a review of the previous literature, or by conducting focus groups (Molin, 2011). According to Louviere and Timmermans (1990a), one should bear in mind the following questions while designing a stated-preference survey: Firstly, which attributes should be ignored in order to achieve the most relevant list of attributes to consumers (respondents)?; Secondly, which attributes should be recombined or grouped to avoid excess attributes?; Thirdly, are the attributes clearly defined and are they suitably represented?

As the starting point in designing our stated-preference surveys, the available literature was first reviewed. A list of potential attributes and their respective levels were compiled and used as a guide during the course of a number of focus groups. Within the literature, Hunt (2001) and Hunt et al. (1996) were two stated-preference studies conducted in Canadian context. A comprehensive review of past revealed and stated-preference literature (see previous section) was also performed, with important factors influencing housing and neighbourhood preference classified into:

1- **Attributes of the dwelling unit:** cost (price, rent or taxes), building size (number of rooms or floor area), age, lot size, building type, quality of construction, state of repair, etc.;
2- **Attributes of the location:** accessibility to workplace, shopping and other non-work activities, as well as schools, public transport quality, etc.;
3- **Attributes of neighbourhood:** density and openness of built form, traffic, noise and air pollution, demographic mix (race and age), crime rate, average income for households in the area, prestige or quality, character and maturity of landscaping, pleasantness and degree of interest, etc.
4- **Characteristics of the household:** income, number of household members, life cycle status and related indicators, car ownership, etc.

The list of attributes initially considered for this stated-preference survey was extensive. On one hand, housing and neighbourhood preference cannot be represented in terms of only a few attributes (Hunt et al., 1996; Kim et al., 2005; Levine & Frank, 2006; Molin & Timmermans, 2003; Walker et al., 2002). On the other hand, the inclusion of more than four to six attributes in a stated-preference survey has been found to render surveys confusing and too taxing for many respondents to process (Caussade, Ortuzar, Rizzi, & Hensher, 2005).

In order to compile a manageable list of attributes, while working on the computer interface of the survey instrument, three focus groups were organized to determine the final list of attributes for the study. In particular, an attempt was made to: (1) narrow down the attributes to the most salient ones relevant to neighbourhood choice; (2) focus more on the attributes of the neighbourhood as opposed to the dwelling; and finally (3) to include attributes that could be represented both visually and textually since the purpose of this study was to compare the two representation methods.

The next section provides the results of the focus groups where the inclusion of potential attributes and the best way to present them (visually or non-visually) were investigated. In summary, based on the three focus group meetings, seven attributes were selected as they appeared to be the most important and easiest to represent using both media (visual and written).

3.1.3. Focus groups

One important step in survey development is the gathering together of individuals representative of the population of interest into groups to discuss the choice that is the subject of interest of the survey. These group discussions are known as focus groups. Focus groups are moderated and led by trained moderators, and are used to ensure that the right alternatives and attributes are included in the survey, and that options making up the choice tasks are described clearly and concisely. Since conducting focus groups costs money, sometimes this step is skipped by researchers and practitioners; this may have implications for the validity of the survey design, and hence the reliability of responses (Ortuzar & Willumsen, 1994).

After having reviewed the literature, we organized three focus groups with 6-8 people to elicit information from prospective respondents about the attributes that were important to them in terms of neighbourhood choice. Each session was audio recorded, and participants were compensated $50 for their time. The focus groups consisted of one hour of discussion regarding the attributes of a neighbourhood which participants considered when deciding where to live, 30 minutes to test the gaming interface, and finally another 30 minutes to discuss how the attributes brought up earlier could be properly visualized, and whether respondents could differentiate between different levels of attributes as they were currently incorporated to the interface.

Since there was no prior 3D simulation of a neighbourhood choice study to base our work on, a great deal of time was spent in finding attributes that could both be presented in visual and text format. One challenge was finding a proper visual representation of attributes, and their levels that would be understood by all respondents in the same way as when they are being presented in text format. During the focus groups, we also took the opportunity to test preliminary versions of the survey instrument. In each focus group, we inquired about the presence of any unwanted elements within the virtual residential environments, tested the ease of using the interface among respondents from different age groups, and defined the ranges of each of the attributes.

These discussions resulted in the selection of seven attributes which were used to develop the pilot study: housing type in the neighbourhood, dwelling price, size of front yard, travel time to work by car, travel time to work by transit, travel time by car to nearest supermarket, travel time by walking to nearest convenience store. Also, it was found in the focus groups that tailoring the average home value attribute presented to respondents was necessary in order to increase the realism of the choice tasks. This was done by first asking respondents to select the price they would expect to pay if they were looking to buy a residence, then, based on the amount that they chose, a customized average home value would appear in the survey. For this question, interviewers explained to respondents that they should answer the question as though they were moving imminently and that they should choose a price that fit their current financial situation.

3.1.4. Pilot study

With the purpose of pre-testing the instrument, the survey was piloted in June 2013. Four trained interviewers set up tables and laptops at the entrance of a number of locations of a Canadian hardware and home improvement retailer, Rona Inc., throughout Montreal region. As an incentive, each respondent was given the chance of winning an iPad tablet by participating in the study. The survey was provided in both English and French. In total 132 respondents were recruited, among which 61 completed the text-only survey and 71 individuals the gaming survey. Figure 1 shows examples of 3D simulations of neighbourhoods in the pilot study, and Figure 2 presents an example of a choice task administered in the pilot study.

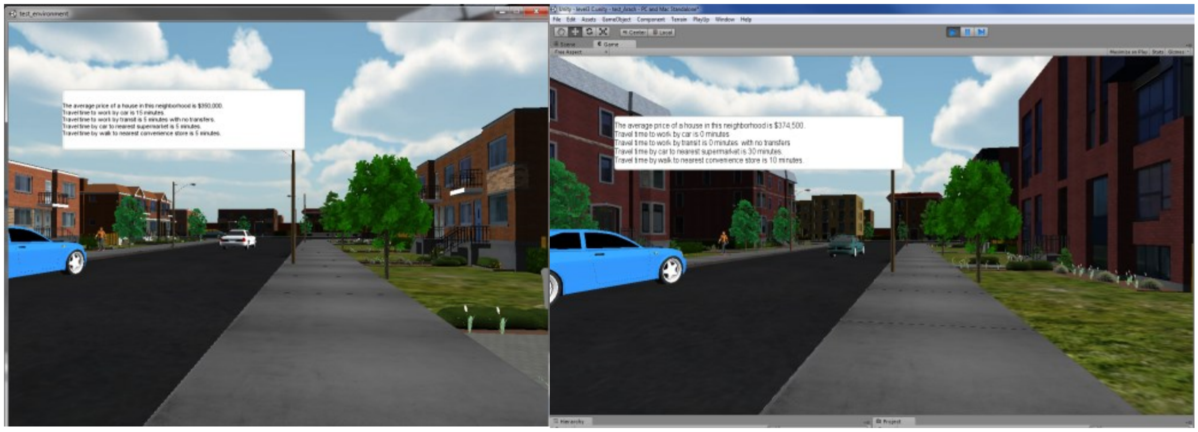


Figure 1 Examples of 3D simulations of neighbourhoods in the gaming platform

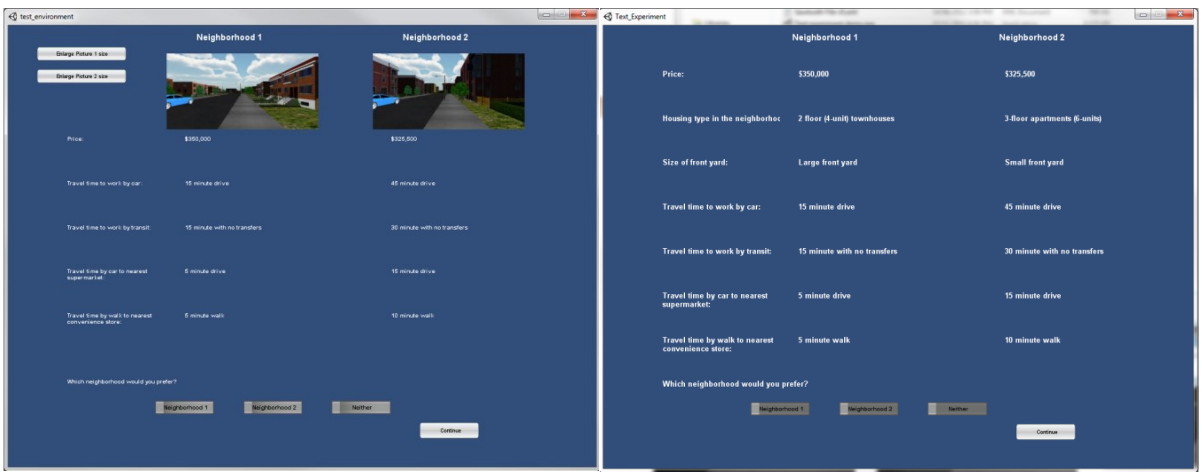


Figure 2 Examples of a choice task of the pilot study in the gaming platform (left) and text-only survey (right)

After reviewing the respondents' comments and analyzing data collected in the pilot testing stage, some necessary changes became clear.

Perhaps the most important change was increasing the number of visual attributes in the final survey instrument. In initial versions of the piloted survey, there were two visual and five textual attributes. Respondents often complained that there was not enough variation in the virtual environments that they were asked to choose between. The two visual attributes (i.e. dwelling type and front yard depth) in our pilot study had three levels. Based on respondent feedback, it was decided that adding an extra level to the attribute of dwelling type (i.e. Triplex) and including an additional attribute, space between buildings, was necessary. At the same time, a written attribute (travel time by car to nearest supermarket) was dropped to avoid survey complexity.

Improving the graphical design of the game interface also seemed essential based on respondents' feedback while administering the survey interfaces. Therefore, the graphical design of text information overlaid on 3D simulations was altered to encourage respondents to read written information while navigating each neighbourhood. At the same time, the configuration of the summary page was redesigned in order to provide larger and clearer images of the two neighbourhoods respondents had just navigated through as part of the choice task. See Figure 5 and Figure 6 for the final interface design.

It was also observed, through the pre-test, that respondents could not differentiate between the levels of front yard depth. In order to make the difference among front yard depth levels more noticeable, the number of levels of this attribute was reduced to two in the final survey. In the pilot survey, the coefficient estimate for the price attribute was found to be insignificant in both surveys and even had a counter intuitive (negative) sign in the text-only survey. It seemed both lower and higher bounds of the price range was deemed acceptable by most of the respondents, and as a result, they did not take price (cost of the dwelling) into account when choosing among neighbourhoods. Upon noticing this, we decided to increase the range of prices presented from 7.5% to 20% plus and minus the base price respondents indicated being comfortable paying for a residence when prompted about this earlier in the survey.

It was also decided to make transit travel time to work partly dependent upon automobile travel time to work, as can be seen in

Table 1. This was done in order to make the transit travel time to work attribute more realistic as participants pointed out that in some alternatives, transit travel times to work were unrealistically competitive with automobile travel times.

Table 1 summarizes the list of attributes and their levels incorporated into the final survey instruments.

Each respondent faced 12 choice tasks in the pilot study. However, the number of choice tasks was reduced to six in the final survey. This was done to decrease the time needed to answer each survey and to avoid compromising data quality due to fatigue. Also, wording of some attributes were refined and ambiguous terms were removed in both English and French versions.

Last but not the least, the sampling methodology also was changed after the pre-test. For the final survey, the interviewers brought laptops to several coffee shops in Montreal and Laval. As we found that people are more approachable at coffee shops whereas, customers of Rona store locations seemed to be in rush to do their errands and were reluctant to spend time answering a survey. We also offered a $5 gift card to each respondent of the same coffee shop where they were being interviewed as an incentive to participate in the survey, and an incentive to the coffee shop to allow us to approach their customers.

Table 1 Summary of Survey Attributes

Attributes		Levels
Dwelling type in neighbourhood		(a) Single Detached Houses (b) 2-storey Townhouses (c) Triplexes* (d) 3-Storey Apartments (6 or 8 units)
Space between buildings		(a) No space (b) 20 feet
Front yard depth		(a) 9 feet (specific to Triplex dwellings) (a) 6 feet deep (for all dwelling types except triplexes) (b) 25 feet deep
Travel time to work by car		(a) 20 minutes (b) 35 minutes (c) 50 minutes
Travel time to work by public transit	**Dependent on time by car**	(a) 18 minutes (b) 25 minutes (when travel time to work by car was 20 minutes) (a) 30 minutes (b) 45 minutes (when travel time to work by car was 35 minutes) (a) 50 minutes (b) 65 minutes (when travel time to work by car was 50 minutes)
Travel time to nearby shops on foot		(a) 5 minutes (b) 15 minutes (c) 25 minutes
Average home value	**Customized**	(a) % 20 below base price (b) Base price (c) % 20 above base price

* Triplexes are flats in 3-storey buildings typical of downtown neighbourhoods in Montreal.

Once the final list of attributes and levels was defined it was necessary to determine how they should be combined and presented. The following section describes the experimental design used in this study.

3.1.5. Experimental Design

SP data is generated by a systematic design and planned variation of attributes and their associated levels. An experimental design is the method of manipulating attributes and their levels. The design deals with the selection of attribute levels characterizing the different choice options across the choice tasks, while ensuring that variables are not correlated. Different design strategies are used based on the purpose and the complexity of the survey instrument, such as full factorial design, fractional factorial design as well as random designs (Louviere et al., 2000).

The simplest design strategy to combine attributes is the full factorial design. It is simply the factorial enumeration of all possible combinations of attribute levels. This design ensures that all attribute effects of interest are truly independent. Moreover, all possible effects associated with analysis of variance can be estimated. A drawback associated with this design is that the size and complexity of stated-preference surveys grow exponentially with increasing the number of attributes and their levels. This makes this design strategy impractical for surveys with great numbers of attributes.

In order to reduce the experiment to a practical size, fractional factorial designs are used. They involve the selection of a subset or sample of the full factorial, bearing in mind particular effects of interest that the research might like to be able to estimate. This results in some loss of statistical information, typically reducing the number of choice tasks in the design with the trade-off of being unable to estimate higher order interaction effects. A random sampling strategy is another approach employed when faced with a complex and large stated-preference survey. If large enough samples from the complete factorial are taken, it would be possible to closely approximate the statistical properties of the full factorial design. The researcher, in this method, needs to select relatively large samples from the full factorial, divide the profile into subsets (blocks) and randomly assign respondent to the blocks.

The field of experimental design, however, is quite complex and as a result it is only mentioned briefly here. Luckily, there are several software packages available to researchers and practitioners to generate designs. In this study, Sawtooth software SSI Web (Sawtooth-Software-Inc, 2013) was used for both surveys. The design algorithm of Sawtooth offers different random design strategies, and is able to generate a fixed set of profiles by drawing from the full factorial design taking into account any prohibitions set by the designers. For the final version of the surveys, 200 different versions of the surveys (blocks), each consisted of 6 choice tasks, were generated with the aim of soliciting 200 persons using each survey instrument.

Among the design strategies available within Sawtooth, the balanced overlap method was used to generate the final design. This approach is between the random and complete enumeration strategies. It allows almost half as much level overlap within the same task as the random method. Although this approach is statistically less efficient than designs with minimal overlap, it has the benefit of encouraging respondents to base their decisions on all attributes of the design, especially in the presence of dominant attributes in the study.

3.1.6. Final survey interface and choice task design

Altogether, the final stated-preference survey consisted of five parts:

1. A consent form explaining the purpose of the study,
2. A question asking the price each respondent would expect to pay for a residence if they were purchasing a home,
3. A tutorial for learning how to navigate the 3D virtual environments (only for gaming platform survey),
4. The stated-preference choice tasks, and
5. Questions on the socio-demographics of the respondents.

The stated-preference survey was built using a gaming engine called Unity (unity3D.com) in order to allow respondents to ‘explore’ a virtual neighbourhood, but also so that all elements of the virtual neighbourhood could be controlled. All parts of the stated-preference survey were developed and administered using a gaming platform, including the generation of the three-dimensional environments; variation of the attributes according to experimental design (and user input); and storing respondent choices and socio-demographic information. The gaming engine was used to produce the text-only version of the survey as well. In the text-only survey, however, the three visual attributes were described using written descriptions. The gaming engine was also able to collect other information, such as the length of time spent on a choice task.

The first step in developing the visual survey was generating 3D models of buildings and neighbourhood features. Google SketchUp-Pro was employed to create the 3D models. Four dwelling types were developed using this software and the other elements were downloaded from Google warehouse database, for instance fences, plants, cars and so on. See Figure 4 for examples of the 3D environments.

While modeling the environments, special care was taken to keep everything consistent across alternatives in order to avoid introducing any unintended variations to the survey. The visual elements of a neighbourhood remained the same in all simulated environments. For example, buildings were placed at a constant setback line, the streets and sidewalks were of identical width and had the same number of trees, parked cars, moving pedestrians as well as the same blue sky.

In each choice task of the gaming platform survey, respondents were required to 'walk' through two different residential streets. Respondents were able to control the direction and speed with which they walked through the neighbourhoods. They were also able 'look around' by panning left, right, up and down. After a few moments of moving through a virtual neighbourhood, written information was superimposed over the visualization. After having walked through the two alternative neighbourhoods of the choice task, they came to a page summarizing what they had seen earlier and were asked to choose their preferred alternative neighbourhood. In the case of the text-only survey, respondents were shown the same information as in the virtual environment survey, only in the form of table. Examples of choice tasks from the stated-preference surveys are shown in Figure 5 and Figure 6.

The next section reports on survey administration, and explains the sampling frame used and sample characteristics.

3.2. Survey Administration

3.2.1. Sampling frame

Once the survey is ready a proper sampling frame should be chosen in order to access to the population of interest. The sampling frame determines the pool of respondents from which a finite sample is drawn to administer the data collection instrument. The objectives of a study define the sampling frame. Simple random samples (SRS) and exogenously stratified random sample (ESRS) are two common sampling strategies.

In the SRS strategy, each individual in the sampling frame has same chance of being selected for the sample while in the ESRS, the sampling population is divided into a few mutually exclusive groups, each representing a stratum of the population. Any characteristics of the participants can be used to create strata, and any individual in the same stratum has the same chance to be chosen (Louviere et al., 2000).

Since this research was not concerned with drawing conclusions about different population groups, and given the difficulty of recruiting people with experience of choosing a residence or neighbourhood, a strict random sampling approach was not adopted. It is worth mentioning that initially, in the pilot study, respondents were pre-screened and only home-owners were invited to complete the survey. Accordingly, a number of Rona Inc., i.e. home improvement and gardening stores, were chosen in the hope of accessing home-owners. This sampling strategy was dropped for administering the final survey instruments and both homeowners and renters were asked to participate. Because excluding renters would have made data collection difficult given the high proportion of rental housing, 43.5% according to Statistics-Canada (2003), in Montreal.

Various sites were chosen in the greater metropolitan region of Montreal to gain access to a cross-section of respondents in terms of demographics and current neighbourhood type. Since the study was primarily concerned with issues related to representational method of stated-preference surveys, a formal representation of the population was not deemed necessary and not sought. Locations were chosen in order to be able to access respondents more likely to have experience with having searched for a residence to buy. Neighbourhoods with a large student population were, for example, avoided. At the same time both homeowners and renters participated in the survey.

In the first round of data collection, the surveys were administered at several coffee shops in June 2013. After preliminary analysis, it was deemed necessary to collect more data for estimation purposes; therefore, another round of data collection was performed in February 2014. The two surveys were administered with laptops. Interviewers were hired and trained in order to help with, and explain, the survey. Interviewers administered both the text-only and gaming platform surveys and they were instructed to behave consistently across respondents of both surveys in order to avoid interviewer or survey version biases. Each of the interviewers had a laptop and was responsible for screening and approaching potential participants. Respondents administered the surveys themselves, although the interviewer was present in case the respondent needed any help or clarification. Both surveys could be done in French or English.

In order to recruit a similar sample for the two surveys, the respondents were randomly and equally divided between the gaming platform and text-only surveys. The text-only survey took each respondent approximately 10 minutes to complete while the gaming platform survey took 20 minutes. As an incentive to participate, each participant was offered a $5 gift card for the coffee shop in which they were being interviewed. In total, 489 respondents completed both surveys. Figure 3 below presents the geographic distribution of respondents.

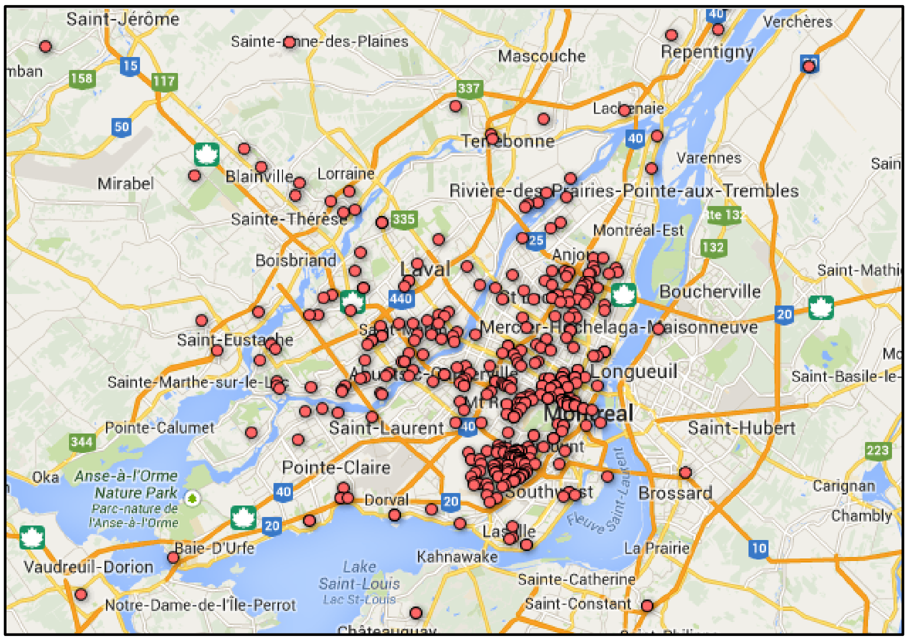


Figure 3 Geographic distribution of respondents

3.2.2. Sample and Population Description

Through inspection of the responses, inconsistent responses and lexicographic responses were excluded. For example, households who always preferred the highest level of a single attribute were excluded. According to Ortuzar and Rodríguez (2002), even if this behaviour effectively corresponds to the respondent preferences, it is not consistent with the compensatory decision structure of the multinomial logit model.

For each representation mode, the final data set after data cleaning had 184 respondents * 6 choice sets, making 1104 observations. These responses were used in the estimation process. Table 2 Descriptive Statistics, presented below, summarizes the characteristics of the sample used to estimated discrete choice models for two representation modes. It reveals that there is a reasonable match between the two sub-samples with some relatively minor differences, suggesting the two sub-samples are comparable. As such, the difference across representation modes may be associated with different methods and not the difference of respondents of the two presentational methods.

Table 2 Descriptive Statistics of the entire sample (before data cleaning)

	Text-only format	Gaming platform
Gender		
Female	48%	41%
Male	52%	59%
Age		
34 and under	29%	34%
35-44	21%	25%
45-54	25%	23%
55-64	13%	16%
65 and above	2%	1%
Employment status		
Full-time	76%	71%
Part-time	7%	10%
Student and employed	9%	10%
Student	3%	4%
Unemployed	5%	4%
Current dwelling tenure type		
Home-owner	52%	55%
Renter	48%	45%
Expected price to pay for residence		
$100,000 - $300,000	62%	63%
$400,000 - $600,000	28%	27%
$700,000 - $900,000	10%	10%
Commute to work		
Car	62%	59%
Public transit	27%	30%
Walk	7%	9%
Bike	4%	2%

The next section describes the data analysis framework used in this study.

3.3. Data analysis

The final step in carrying out a stated-preference study is conducting model estimation and analyzing collected data. Random utility models are commonly used to describe decision makers' choices between alternatives. This section explains the modeling framework used in this study. In this modelling approach, the alternatives represent competing products, policies or any options among which choices should be made.

3.3.1. Random Utility Theory

The use of random utility theory (RUM) is a well-established approach for estimating discrete choice models. Discrete choice models are usually derived under an assumption of utility-maximizing behaviour by the decision maker. According to this theory, it is assumed that a decision maker facing a choice among a set of alternatives obtains utility from each alternative and chooses the alternative that provides the greatest utility. The behavioural model thus becomes:

Equation 1 Utility maximization in random utility theory

$$U_{ni} > U_{nj} \; \forall j \neq i$$

This utility is unknown to the researcher while some attributes of the choice alternatives and decision makers are observed by the researcher. The RUM therefore associates an individual's choices with the observed characteristics (V_{ni}), or systematic utility, and an error term (ε_{ni}) that captures the unobserved factors influencing choice (Train, 2009).

Equation 2 Decomposition of Random Utility

$$U_{ni} = V_{ni} + \varepsilon_{ni} \qquad \forall i$$

Systematic utility is typically considered to be a linear combination of alternatives and respondent characteristics(x_i):

Equation 3 Decomposition of Random Utility

$$U_{ni} = \alpha_{ni} + \beta x_{ni} + \varepsilon_{ni} \qquad \forall i = 1, \dots, j$$

where, α_{ni} is a constant, x_{ni} is a vector of attributes of the alternatives and respondent socio-economic characteristics and ε_{ni} is a random error term.

Researchers consider the error term ($\varepsilon_{ni} \forall i$) to be random. Assumptions about the distribution of the error term define the resulting discrete choice models.

3.3.1.1. Logit model

The logit is the simplest and most commonly used discrete choice model. Under the assumption that the random terms in a decision maker's utility are independently, and identically extreme value distributed (IID), the following closed-form expression for the logit choice probability is derived:

Equation 4 Logit model

$$P_{ni} = \frac{e^{\beta x_{ni}}}{\sum_j e^{\beta x_{nj}}}$$

where, x_{ni} is a vector of attributes of the alternatives and respondent socio-economic characteristics at parameter β

3.3.1.2. Mixed logit model

The majority of empirical studies do not go further than using multinomial logit models. However, more recently, some researchers have used more advanced models in an attempt to increase the behavioural realism of discrete-choice models. Most of the efforts are devoted to relaxing assumption associated with IID (independent and identically distributed) error term to an extent that is behaviourally more enriching, computationally tractable and practical. One of these advanced methods is the mixed logit (ML) model and it is a generalization of the MNL model (Louviere et al., 2000).

A mixed-logit modeling approach is highly flexible in approximating choice models with repeated responses. It overcomes three limitations of the standard logit model by permitting for random taste variation, unrestricted substitution patterns, and correlation in unobserved factors over time.

Mixed logit probabilities are integrals of standard logit probabilities over a density of parameters.

Equation 5 Mixed-logit model

$$P_{ni} = L_{ni}(\beta)\, f(\beta)\, d\beta$$

Here, $L_{ni}(\beta)$ is the logit probability calculated at parameter β and $f(\beta)$ is a density function determined by the researcher. The choice probability integral has no closed-form solution and the parameters are estimated through simulation. However, if the utility is linear in β, the mixed logit probability takes the usual logit form:

Equation 6 Mixed logit model probability

$$P_{ni} = \int \frac{e^{\beta x_{ni}}}{\sum_{j=1}^{J} e^{\mu \beta x_{nj}}}\, f(\beta) d\beta$$

where, x_{ni} is a vector of attributes of the alternatives and respondent socio-economic characteristics at parameter β and $f(\beta)$ is a density function determined by the researcher.

Distributions are arbitrary approximations to the real behavioural profiles and researchers select the ones that best match the behaviour under study. The most commonly used distributions in the literature are the normal and lognormal distributions. There are, however, other distributions used by researchers, such as uniform, triangular, etc. (Train, 2009).

SP data sets are repeated responses (panel data) as each respondent faces a series of choice scenarios in a survey. Therefore, a mixed logit modeling approach is appropriate since it treats the coefficients that enter utility function as varying over individuals but being constant over choice scenarios for each person (Train, 2009).

3.3.3.3. Scale Parameter

In all choice models derived from random utility theory, the choice probabilities are result of certain assumption about the distribution of random part of the utility function. The logit formula, for example, is derived by assuming that the unobserved factors are distributed IID extreme value with variance of $\pi^2/6$. The full expression of the MNL model probability therefore becomes:

Equation 7 Logit model with scale parameter

$$P_{ni} = \frac{e^{\mu \beta x_{ni}}}{\sum_{j=1}^{J} e^{\mu \beta x_{nj}}}$$

Here, μ is the scale parameter, and it is the inverse of the variance of the error term. By definition, and by convention, the mean of error values across observations equals zero and the scale of the error term refers to the size of the variance across observations (Arentze et al., 2003). The scale differences therefore could be used to investigate the extent of inconsistent responses by respondents. However, As discussed in Ben-Akiva and Lerman (1985), the estimated utility parameters are confounded with the scale parameter so that the estimated coefficients are actually equal to μβ. For this reason, the scale parameter is not identifiable for a single model and is set to one for any given model (Hensher, Louviere, & Swait, 1998; Swait & Louviere, 1993).

Different data collection methods, however, may be characterized by different error variance, and thereby might be influenced in different proportions by unobserved influences on the choices being analyzed. When considering datasets coming from different populations or data collection methods, the scale parameter of other datasets can be calculated relative to the reference dataset (Swait & Louviere, 1993).

Assuming K data sources that are combined, it is possible to assume the coefficients of the K data sources to be the same ($\beta_{1=}$ ... = $\beta_{k=}$ β) and to fix the scale parameter for one of the data sources to one. If this done, it is then possible to estimate the relative size of the scale parameters of the remaining K-1 data sources. It is then possible to test whether the K-1 scale parameters are the same as the base data source. Knowing the relative size of the scale parameters makes it possible to evaluate the relative importance of unobserved factors in the choices made in the different data sources (Swait & Louviere, 1993). If there is no difference in the relative size of the scale parameters, coefficients for the different data sources can be compared directly.

3.3.2. Assessment of Survey Differences

The validity of stated-preference data collected through two interfaces cannot be tested separately. However, the results of models estimated using data obtained from two representation methods can (and have been) compared in terms of internal validity with respect to: (1) goodness-of-fit of the models (Arentze et al., 2003; Jansen et al., 2009; Orzechowski et al., 2005; Vriens et al., 1998) , (2) the presence of scale differences across models (Arentze et al., 2003; Orzechowski et al., 2005) as well as (3) relative attribute size (Arentze et al., 2003; Jansen et al., 2009; Orzechowski et al., 2005; Vriens et al., 1998), (4) respondent heterogeneity (Vriens et al., 1998), (5) the number of significant attribute main effects (Arentze et al., 2003; Holbrook & Moore, 1981; Jansen et al., 2009; Orzechowski et al., 2005; Vriens et al., 1998), and finally the external validity could be determined by investigating the extent the estimated models could predict a revealed preference data base and/or a hold out sample (Levine & Frank, 2006; Orzechowski et al., 2005; Train, 2009).

As a starting point, separate MNL models may be estimated along with a combined MNL model with scale parameter. Then, mixed logit models were used to investigate the extent of response heterogeneity in both data sets. There are various software packages to simultaneously estimate both utility coefficients and the scale parameter, such as Biogeme (Bierlaire, 2003).

4. INTRODUCTION TO MANUSCRIPTS (CONTRIBUTION BY AUTHORS)

The following chapter is a co-authored paper, entitled “Gaming platform vs. traditional text-only stated-preference survey of neighbourhood choice”, submitted to the Journal of Housing and the Built Environment. The paper was part of a research project led by Dr. Zachary Patterson, first author on the paper, funded by Canada Research Chairs program and the Canadian Foundation for Innovation. The objective was to provide insight into the influence of representation method on the reliability and the nature of result of a stated-preference survey of neighbourhood choice.

Preliminary results of this study were presented at the Transportation Research Board’s 93rd annual meeting in Washington, DC, in January of 2013.

The final manuscript presented here was the result of improvement of the first paper and the application of more advanced modelling approaches. I was responsible for the literature review, setting up the survey instruments, administering the surveys, analysis of data, and writing up the results while Sydney Swaine-Simon, a computer science student helped on programming the two versions of the survey. Dr. Ali Rezaei was extensively involved in providing guidance in the design of the survey interfaces and helped in estimating statistical modelling used in the paper. Dr. John Zacharias, the last author, helped in the preliminary stages of the research by providing insights into the urban design elements of the 3D simulations of residential neighbourhoods in our survey interface.

This paper is built on the comprehensive literature review and methodology described in previous sections. This section contains the final paper submitted which is followed by overall conclusions of this study and possible future avenues of research in section 6.

5. GAMING PLATFORM VS. TRADITIONAL TEXT-ONLY SP SURVEY OF NEIGHBOURHOOD CHOICE

Zachary Patterson (corresponding author)
Department of Geography, Planning and Environment, Concordia University
1455 de Maisonneuve W., H 1255-15 (Hall Building)
Montreal (QC), Canada   H3G 1M8
Tel: (514) 848-2424 ext. 3492
E-mail: zachary.patterson@concordia.ca

Javad Mostofi Darbani
Department of Geography, Planning and Environment
1455 de Maisonneuve W. H 1255-15 (Hall Building)
Montréal (QC), Canada H3G 1M8
Tel.: (514) 848-2424 ext. 3310
E-mail: jj.mostofi@gmail.com

Ali Rezaei
Department of Geography, Planning and Environment
1455 de Maisonneuve W. H 1255-15 (Hall Building)
Montréal (QC), Canada H3G 1M8
Tel.: (514) 848-2424 ext. 3310
E-mail: a.rezaaei@gmail.com

John Zacharias
College of Architecture and Landscape
Peking University
Beijing 200080 China

ABSTRACT

Stated Preference (SP) surveys on housing and neighbourhood choice have traditionally been administered in text-only format. It has been argued alternately that the visual presentation of attributes can dominate in SP surveys or that they can be used to help improve the realism of choice tasks, and/or increase the number of attributes that can be included in the surveys. A few studies have tested the difference between multimedia and text-only SP surveys of housing choice. While these studies have been informative, they have drawn conflicting conclusions and have been based on small sample sizes. The research presented here sought to take advantage of the capabilities available in gaming engines to compare the results of SP surveys of neighbourhood choice administered either on a gaming platform, or as a traditional text-only survey. A sample of 368 (184 for each survey) respondents was used to compare neighbourhood choice models from data administered with the two different survey methodologies. We found that while both surveys result in similar models, the model estimated on data obtained from the gaming platform shows slightly better performance. Providing 3D simulations appears to better focus respondent attention. At the same time, it has the drawback that it may decrease the generalizability of some visual parameter estimates because of the need to provide precise visual representations of their characteristics.

Keywords: Stated-Preference surveys, neighbourhood choice, attribute visualization, gaming engines

5.1 INTRODUCTION

The question of how best to provide information in Stated-Preference (SP) surveys has received a fair bit of attention in the literature. While some have argued for the use of visualization techniques to present certain types of information (Dijkstra et al., 1996; Jansen et al., 2009; Levine & Frank, 2006; Morrow-Jones et al., 2004; Orzechowski et al., 2005; Rid & Profeta, 2011), controversy remains on whether one should, and how best to use visualization techniques, and even whether there are any advantages to providing visual information (Caussade et al., 2005; Dijkstra et al., 1996; Hensher et al., 1998; Hunt et al., 1996; Louviere et al., 2000; Schall, Schöning, Paelke, & Gartner, 2011; Wittink et al., 1994). This study reports on research that seeks to address these questions in the context of neighbourhood choice by administering two SP surveys, one through a gaming platform and the other as a traditional text-only survey.

In the gaming survey, respondents were able to navigate three-dimensional simulated neighbourhoods in which they received supplementary textual information as they viewed alternative virtual neighbourhoods. By contrast, the text-only survey employed traditional means of representation so that all attributes were presented as written descriptions only.

Ensuring respondents can process all attributes presented to them and ensuring realism are two important, yet often conflicting goals in the design of SP surveys (Hensher et al., 1998). Neighbourhood choice involves many potential attributes, and therefore cannot be realistically represented with too few (Molin & Timmermans, 2003). The goal of this study was to see whether there were advantages to presenting information using a combination of visual and written attributes in a gaming platform, as opposed to a traditional text-only approach. In particular, we sought to examine whether the two presentation methodologies produced different statistical results in the analysis of the SP data.

Both surveys were programmed using the gaming engine Unity, which allowed for the creation of 3-dimensional virtual environments for the gaming version of the survey, as well as a text-only version. In addition to being a good medium for the development of virtual neighbourhood environments, it also allowed great flexibility in the design and customization of attribute values, such as average home value and travel time, to respondent input.

The paper starts with a review of existing relevant literature and is followed by a section on methodology that explains survey development and interface design, as well as how the survey was administered. The data are then analyzed and the modeling results presented. After discussion of the modeling results, a final section discusses the contribution of the results and suggests some avenues for future research.

5.2 LITERATURE REVIEW

In the past, there has been a significant amount of research to evaluate the use of visual techniques to provide information about attributes in stated-preference surveys. A number of studies in SP research have discussed benefits of visualizing attributes. First, visualization may enhance respondent understanding and decrease the ambiguity of choice tasks which leads to more heterogeneity in responses. Vriens et al. (1998) found higher response heterogeneity when images were used in stated-preference surveys. That is, they were able to have more segments for one of the visual attributes in the latent class model they developed, concluding that this may be the result of a more informed evaluation by the respondents in the visual representation mode.

Second, it has been suggested in the literature that more attributes can be included in choice tasks while reducing the risk of information overload, when some attributes are presented visually (Wittink et al., 1994). In line with this, Holbrook and Moore (1981) found more significant variables in their SP surveys thanks to the inclusion of drawings in their surveys of sweater choice. This, however, is not a universal finding; Vriens et al. (1998) for example found no difference in the number of significant attributes in a similar study related to car stereo design.

Third, pictorial representations of some attributes have been expected to improve the realism of tasks since they can better mimic the actual inspection of a product in the market-place and therefore that respondents would be more likely to make the same choices in the surveys as in real-world situations (Dijkstra et al., 1996; Green & Srinivasan, 1978; Vriens et al., 1998; Wittink et al., 1994).

Others, on the other hand, have emphasized the drawbacks of using visual information in SP surveys. Arentze et al. (2003) investigated the effect of representation mode as well as task complexity in terms of the number of attributes, alternatives and choice sets as well as literacy of the respondents in a stated-preference survey concerned with transport mode choice. They found no significant increase in error variance of the model (scale parameter difference) or the attribute weights across the two presentation methods (pictorial representation and text-only description). However, they did find substantial difference in the error variance when the number of attributes was increased from three to five. As a result, the attribute weights were changed after correcting for the scale difference. They suggest it was not worth the extra effort to develop visual materials for their survey, as no difference was observed between the presentation methods.

Stated Preference (SP) surveys in the domain of housing and neighbourhood choice have typically been presented in the traditional text-only format (Cooper et al., 2001; Hunt, 2001; Hunt et al., 1996; Kim, 2006; Kim et al., 2005; Louviere & Timmermans, 1990a; Ortuzar et al., 2000; Senior et al., 2006; Walker et al., 2002; Wang & Li, 2004). There have been, however, a few studies to apply multimedia approaches (i.e. a combination of written and visual representations) to these SP surveys (Jansen et al., 2009; Levine & Frank, 2006; Morrow-Jones et al., 2004; Orzechowski et al., 2005). Some of the few studies that have adopted a multimedia approach and attempted to compare the results obtained across representation modes (i.e. Jansen et al. (2009) and Orzechowski et al. (2005)), have identified important implications for the estimated parameters of the resulting statistical choice models, and have noted both benefits and drawbacks of these approaches.

With respect to the traditional approach to SP surveys in housing and neighbourhood choice, Hunt (2001) is a good example, and is perhaps the most comprehensive in a Canadian context. This research examined the impacts of transportation-related factors on the attractiveness of residential locations in Edmonton, Canada. This study can be considered a further development of his earlier work (Hunt et al., 1996). The SP study combines nineteen variables into each housing alternative and holds four of the attributes constant across alternatives to decrease the task complexity. The study uses a large sample size with the total number of respondents consisting of 1277 interviews. In order to establish a consistent understanding of the attributes, supplemental materials were provided in the form of pieces of paper with bullet points, and in some cases photographs.

With respect to the use of visual information in housing and neighbourhood SP surveys, two main studies have been found. Jansen et al. (2009) conducted two studies to examine the impact of including images along with text in SP surveys of housing choice. In the first study, three different versions of their survey were used: one text-only version, another with text and black and white drawings of the alternative houses, and a third with color photos. The surveys had 5 attributes and were completed by the same 28 respondents. In their second study, they compared a text survey with photos available on request and another with text and photo montages. Both surveys consisted of 13 attributes, 7 attributes related to dwelling characteristics and six for characteristics of the dwelling environment. 48 respondents filled out the text survey and 59 completed the survey with text and photo montages. Jansen et al. (2009) found differences in estimated parameters between the SP surveys that they suggest could be the result of unsystematic variation and unwanted details in images presented in the surveys with images. They also found that visually presented attributes gained importance compared to the same attributes when they were presented in text. It is argued that this may be because respondents have an inclination to process visual elements more readily than written text.

Orzechowski et al. (2005) conducted another study relating to housing choice that compared a multimedia presentation methodology (Virtual Reality) with a traditional text-only survey. The multimedia version of the survey allowed respondents to change their perspective on 3D architectural plans of alternative houses with different layouts. Four attributes were used in these surveys, among which price was the only attribute presented in a text format in the multimedia survey. Their research found no evidence of significant differences in terms of internal and external validity between the two methods. At the same time, they found coefficient estimates of the price attribute to be more believable in models developed from data in the text-only survey (the estimate of the highest price level was insignificant, and the magnitudes of coefficients for the other price levels were inconsistent in the visual survey). The authors suggest that this is because the visual attributes dominate when in the presence of attributes presented with text. Finally, the authors expected less variance of the error term in models developed using data collected with the multimedia representation mode since the graphical presentation of attribute levels was assumed to make the attribute interpretation easier. This was not borne out in the results, however–the visual survey models did have a lower error variance but the difference in error variance was not statistically significant. This result may have been due to the fact that there were relatively few respondents to the survey with 35 respondents for the text-only survey and 29 for the multimedia survey.

To summarize, existing literature on SP surveys of housing and neighbourhood choice has, for the most part, been based on traditional text-only surveys. Few studies have attempted to use visual techniques to present choice tasks to respondents and compare the differences obtained between representation modes. Despite the benefits pointed out in the literature for the visual presentation of attributes, past studies have: found that visual attributes take on more importance, or reduce the importance of variables presented as text; had trouble ensuring systematic control of visual attributes (i.e. with Jansen because of the use of pictures, renderings and collages of pictures); and had small numbers of attributes describing alternatives. Moreover, the methodological studies comparing different representation modes have had small sample sizes.

In the present research we sought to fill some of these gaps by using a gaming engine to develop and administer an interactive, customized, SP survey of neighbourhood choice with a relatively rich set of attributes and relatively large sample sizes. The use of the gaming engine allowed complete control over the simulated environment to avoid the potential for unsystematic variation negatively affecting results.

5.3 METHODOLOGY

The list of attributes initially considered for this SP surveys was extensive. As Molin and Timmermans (2003) point out, housing and residential choice cannot be represented in terms of only a few variables. On the other hand, the inclusion of more than four to six attributes in an SP survey has been found to render surveys confusing and too much for respondents to process (Caussade et al., 2005).

In order to compile a manageable list of attributes, an attempt was made to: (1) narrow down the attributes to the most salient ones relevant to neighbourhood choice; (2) focus more on the attributes of the neighbourhood as opposed to houses; and finally (3) to include attributes that could be represented both visually and textually since the purpose of this study was to compare both presentation methodologies.

Table 3 Summary of Survey Attributes

Attributes		Levels
Dwelling type in neighbourhood		(a) Single Detached Houses (b) 2-storey Townhouses (c) Triplexes* (d) 3-Storey Apartments (6 or 8 units)
Space between buildings		(a) No space (b) 20 feet
Front yard depth		(a) 9 feet (specific to Triplex dwellings) (a) 6 feet deep (for all dwelling types except triplexes) (b) 25 feet deep
Travel time to work by car		(a) 20 minutes (b) 35 minutes (c) 50 minutes
Travel time to work by public transit	**Dependent on time by car**	(a) 18 minutes (b) 25 minutes (when travel time to work by car was 20 minutes) (a) 30 minutes (b) 45 minutes (when travel time to work by car was 35 minutes) (a) 50 minutes (b) 65 minutes (when travel time to work by car was 50 minutes)
Travel time to nearby shops on foot		(a) 5 minutes (b) 15 minutes (c) 25 minutes
Average home value	**Customized**	(a) % 20 below base price (b) Base price (c) % 20 above base price

* Triplexes are flats in 3 storey buildings typical of downtown neighbourhoods in Montreal.

After having reviewed the literature, we organized three focus groups with 6-8 people to elicit information from prospective respondents about the attributes that were important to them in terms of neighbourhood choice. During the focus groups, we also took the opportunity to test preliminary (but evolving) versions of the survey instruments. As such, in each focus group we inquired about the presence of any unwanted elements within the virtual residential environments, tested the ease of using the interface for different age groups, and defined the ranges of each of the attributes. These discussions resulted in the selection of seven attributes for the surveys that are presented in Table 3.

In order to increase the realism of the choice tasks, average home value attribute was customized to respondent input. First, respondents were asked to select the price they would expect to pay if they were looking to buy a residence. The amount that they chose was then used to customize the average home value that appeared in the survey. For this question, interviewers explained to respondents that they should answer the question as though they were moving imminently and that they should choose a price that fit their current financial situation.

While not customized to respondent input, transit travel time to work was partly dependent upon automobile travel time to work, as can be seen in Table 3. This was done in order to make the transit travel time to work attribute more realistic - in early versions of the survey, participants pointed out that in some alternatives, transit travel times to work were unrealistically competitive with automobile travel times.

In initial versions of the survey, there were only two visual and five written attributes. The visual attributes consisted of three levels of dwelling type and two levels of front yard depth. Respondents often complained that there was not enough variation in the virtual environments that they were asked to choose between. As a result, a fourth level was added to the attribute of dwelling type (Triplex), and space between buildings attribute was added as an attribute.

5.3.1 Survey Interface Development and Description

Since the purpose of the study was to compare the statistical choice model results obtained from two SP surveys with different representation methodologies, two versions of the same SP survey had to be developed. The first version was a conventional text-only survey where all attribute information was presented as text. The second was a multimedia version administered as a gaming platform, where respondents received supplementary textual information as they navigated a residential street in a 3D virtual environment.

The SP survey was built using the Unity (unity3D.com) gaming engine. Gaming engines are software used to develop video games. A gaming engine was used first in order to allow respondents to ‘explore’ a virtual neighbourhood, but also so that all elements of the virtual neighbourhood could be controlled. Initial research considered using Google Streetview as a way to navigate neighbourhoods, but it was too difficult to control for the many differing characteristics of real neighbourhoods.

All parts of the SP survey were developed and administered using this gaming platform, including the generation of the three-dimensional environments; variation of the attributes according to experimental design (and user input); and storing respondent choices and socio-demographic information. The gaming engine was used to produce the text-only version of the survey as well. In the text-only survey, however, the three visual attributes were described using written descriptions. The gaming engine was also able to collect other information, such as the length of time spent on a choice task.

The first step in developing the visual survey was generating 3D models of buildings and neighbourhood features. Google SketchUp-Pro was employed to create the 3D models. Four dwelling types were developed using this software and the other elements were downloaded from a Google warehouse database, for instance fences, plants, cars and so on. See Figure 4 for examples of the 3D environments.

Figure 4 Examples of levels of housing type in gaming platform

While modeling the environments, special care was taken to keep everything consistent across alternatives in order to avoid introducing any unintended variations to the survey. The visual elements of a neighbourhood remained the same in all simulated environments. For example, buildings were placed at a constant setback line, the streets and sidewalks were of identical width and had the same number of trees, parked cars, moving pedestrians as well as the same blue sky.

In each choice task of the gaming platform survey, respondents were required to 'walk' through two different residential streets. Respondents were able to control the direction and speed with which they walked through the neighbourhoods. They were also able 'look around' by panning left, right, up and down. After a few moments of moving through a virtual neighbourhood, written information was superimposed over the visualization. After having walked through the two alternative neighbourhoods of the choice task, they came to a page summarizing what they had seen earlier and were asked to choose their preferred alternative neighbourhood. In the case of the text-only survey, respondents were shown the same information as in the virtual environment survey, only in the form of table. Examples of choice tasks from the SP surveys are shown in Figure 5 and Figure 6.

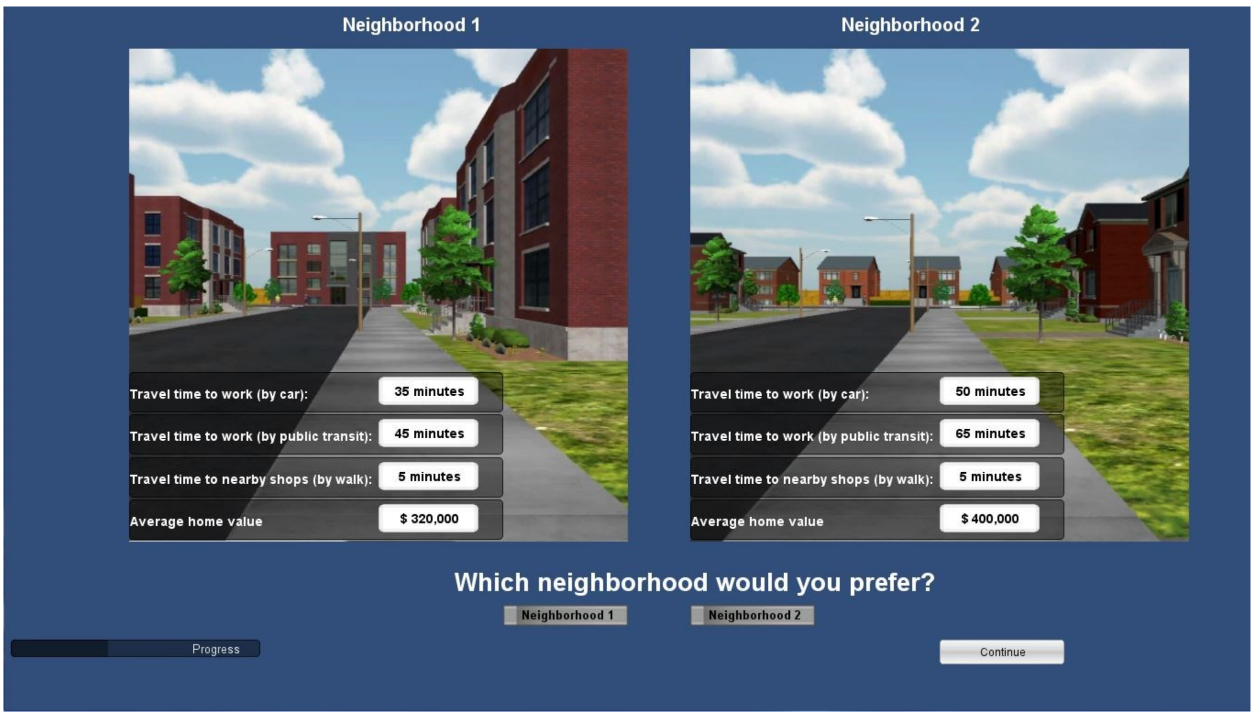


Figure 5 Example of gaming platform choice task summary

Figure 6 Example of text-only choice task

Altogether, the final SP survey consisted of five parts:

1. A consent form explaining the purpose of the study,
2. A question asking the price each respondent would expect to pay for a residence if they were purchasing a home,
3. A tutorial for learning how to navigate the 3D virtual environments (only for gaming platform survey),
4. The SP choice tasks, and
5. Questions on the socio-demographics of the respondents.

5.3.2 Experimental Design

The experimental design used in this study was generated with Sawtooth software SSI Web (Sawtooth-Software-Inc, 2013). The same experimental design was used for both surveys. The design algorithm of Sawtooth software offers different random design strategies and is able to generate a fixed set of profiles by drawing from the full factorial design taking into account any prohibitions set by the designers.

The balanced overlap method was used to generate the final design. This approach is between the random and complete enumeration strategies. It allows almost half as much level overlap within the same task as the random method. Although this approach is statistically less efficient than designs with minimal overlap, it has the benefit of encouraging respondents to base their decision on all attributes of the design, especially in the presence of dominant attributes in the study.

5.3.3 Survey Administration

In the first round of data collection, the surveys were administered at several coffee shops in June 2013. Later, in February 2014, surveys were again administered to collect additional data. Various sites were chosen in the greater metropolitan region of Montreal to gain access to a varied cross-section of respondents in terms of demographics and current neighbourhood type. Since the study was primarily concerned with issues related to presentation methodologies of SP surveys, a formal representation of the population was not deemed necessary and was not sought. Locations were chosen in order to be able to access respondents more likely to have experience with having searched for a residence to buy. Neighbourhoods with a large student population were, for example, avoided. At the same time both homeowners and renters participated in the survey. Excluding renters would have made data collection even more difficult given the high proportion of rental housing, 43.5% according to Statistics-Canada (2003) in Montreal.

The two surveys were administered with laptops. Interviewers were hired and trained in order to help with, and explain, the survey. Interviewers administered both the text-only and gaming platform surveys and they were instructed to behave consistently across respondents of both surveys in order to avoid interviewer or survey version biases. Each of the interviewers had a laptop and was responsible for screening and approaching potential respondents. Respondents administered the surveys themselves, although the interviewer was present in case the respondent needed any help or clarification. Both surveys could be done in French or English since the majority of the population of Montreal is francophone.

In order to recruit a similar sample for the two surveys, the respondents were randomly and equally divided between the gaming platform and text-only surveys. The text-only survey took each respondent approximately 10 minutes to complete while the gaming platform survey took 20 minutes. As an incentive to participate, each participant was offered a $5 gift card for the coffee shop in which they were being interviewed. For each representation method, the final data set after data cleaning had six sets * 184 respondents, making 1104 observations.

Table 4 shows that there is a reasonable match between the two sub-samples with some relatively minor differences, suggesting the two sub-samples are comparable. As such, the difference across representation modes may be associated with different methods and not the difference of respondents of the two presentational methods.

Table 4 Summary of Descriptive Statistics

	Text-only format	Gaming platform
Gender		
Female	48%	41%
Male	52%	59%
Age		
34 and under	29%	34%
35-44	21%	25%
45-54	25%	23%
55-64	13%	16%
65 and above	2%	1%
Current dwelling tenure type		
Home-owner	52%	55%
Renter	48%	45%
Expected price to pay for residence		
$100,000 - $300,000	62%	63%
$400,000 - $600,000	28%	27%
$700,000 - $900,000	10%	10%

5.4 DATA ANALYSIS AND RESULTS

As is common practice in the analysis of SP surveys, discrete choice models based on random utility theory were estimated for data from each of the survey versions. According to random utility theory, a decision maker (n) choosing between different available alternatives (i) chooses the one that provides the highest utility. The indirect utility function is written as:

$$U_{ni} = \alpha_{ni} + \beta x_{ni} + \varepsilon_{ni} \qquad \forall i = 1, \dots, j$$

where α_{ni} is a constant, x_{ni} is a vector of attributes of the alternatives and respondents socio-economic characteristics and ε_{ni} is a random error term. If the error term is identically and independently distributed (IDD) type I-extreme value with scale parameter μ the probability of individual n choosing alternative takes the well-known form of the multinomial logit:

$$P_{ni} = \frac{e^{\mu \beta x_{ni}}}{\sum_{j=1}^{J} e^{\mu \beta x_{nj}}}$$

Here, μ is the scale parameter, and is the inverse of the variance of the error term. The scale parameter is not identifiable for a single model and is set to one for any given model. However, if combining K data sources, it is then possible to fix the scale parameter for one of the data sources to one and estimate the relative size of the scale parameters of the remaining K-1 data sources (Hensher, Rose, & Greene, 2005; Swait & Louviere, 1993).

Given that our respondents faced six different choice scenarios the data sets become a panel of responses. A mixed-logit modeling approach is highly flexible in approximating choice models with repeated responses (Train, 2009). We assume the coefficient estimates vary over respondents but are constant for each individual. Accordingly, the probability that a given individual chooses any alternative becomes:

$$P_{ni} = \int \frac{e^{\mu \beta x_{ni}}}{\sum_{j=1}^{J} e^{\mu \beta x_{nj}}} f(\beta) d\beta$$

Here, $f(\beta)$ denotes a distribution density. The researcher specifies a distribution for the coefficients. The choice probability integral has no closed-form solution and the parameters are estimated with simulation.

5.4.1 Modeling Results

In this section we compare the results of models using data obtained from the two representation modes with respect to: (1) overall goodness of fit of the models, (2) coefficient values, (3) the significance of model coefficients and (4) the ability to account for respondent heterogeneity. First, separate multinomial logit (MNL) models were estimated. Then, mixed-logit (ML) models were used to investigate the extent of response heterogeneity in both data sets. The comparison of the representation modes is discussed based on the results of ML models since they had higher final log-likelihoods, reflecting the importance of taking into account respondent heterogeneity. Ideally, the models would have been estimated with a random sample of 80% of respondents and models validated on the remaining 20% of respondents. Given the difficulty of recruiting respondents we did not have enough data to make this approach possible. Table 5 summarizes the results of ML models.

Before comparing the models derived from the two representational modes, a pooled model with scale parameter was estimated in a ML model using Biogeme (Bierlaire, 2003). After allowing the scale parameter of the gaming platform survey to vary relative to the text-only survey, no significant difference in the scale parameters of the two data sources was observed. Thus the coefficients of the two models could be compared directly and there was no need for correcting for scale difference.

The ML models are estimated with log-normal distributions for the relevant travel time attributes, and “average home value” to ensure their signs were non-positive over their range. The normal distribution was determined appropriate for the rest of coefficients as there was no prior assumption concerning their sign. 5000 Halton draws were used in the ML estimations.

The log-likelihood of the gaming platform model is slightly higher (less negative) than the text-only model, with an adjusted rho-square of 0.243 for the gaming model and 0.229 for the text-only model. This indicates slightly better model performance for gaming platform. While insignificant in both surveys, alternative specific constants (indicating whether the alternative was on the left or the right) were included in the models to account for unobserved factors not captured by the other variables.

Table 5 Mixed-Logit Model Results

Parameter Name		Mixed logit model			
		Text-only format		Gaming platform	
		Estimate	P-value	Estimate	P-value
Left alternative constant		--	--	--	--
Right alternative constant		-0.054	0.527	0.152	0.110
Dwelling type	Detached houses	Omitted		Omitted	
	Townhouse	-0.468	0.008	-0.364	0.069
	Triplex	-0.967	0.000	-1.349	0.000
	Triplex (Urban core residents)	--	--	1.283	0.002
	Apartment	-1.395	0.000	-1.023	0.000
Space between buildings (in feet)		0.032	0.000	0.030	0.001
Front yard depth (in feet)	Whole sample	0.015	0.005	--	--
	Urban core residents	--	--	-0.031	0.022
Travel time to work by car (in minutes)		-0.074	0.000	-0.072	0.000
Travel time to work by transit (in minutes)	Whole sample	-0.012	0.098	-0.019	0.023
	Non drivers	--	--	-0.023	0.017
Travel time to nearby shops on foot (in minutes)		-0.033	0.000	-0.060	0.000
Average home value (in thousands CDN)	Base price 300K and less	-0.007	0.000	-0.014	0.000
	Base price 400K and more	-0.001	0.322	-0.004	0.000
Standard deviation for random parameters					
Triplex		--	--	0.998	0.002
Apartment		1.140	0.000	1.271	0.000
Space between buildings		0.025	0.028	0.043	0.000
Travel time to work by car		0.063	0.001	0.142	0.037
Number of observations (choice tasks)		1104		1104	
Null log-likelihood		-765.234		-765.234	
Final log-likelihood		-576.069		-563.294	
adjusted Rho-square		0.229		0.243	
Likelihood ratio test		25.65		42.22	

The coefficients for dwelling type are expressed in relation to the category of single detached houses. For both townhouses and apartments, the estimated coefficients are negative suggesting respondents would prefer single detached houses to these dwelling types, which is consistent with our expectations and therefore right-sided. At the same time, the magnitudes of these coefficients are smaller, in absolute terms, in the gaming survey. The townhouse with a value of -0.364 in the gaming model is less than the -0.468 in the text-only model. Similarly, the apartment coefficient (-1.023) is smaller than in the text-only survey. This suggests that respondents were less averse to townhouses and apartments in the gaming model. This is a notable result that will be discussed below. Results for the triplex coefficient are slightly different. First, in the gaming platform model it was possible to segment the triplex coefficient according to whether respondents lived in central Montreal neighbourhoods – neighbourhoods with high concentrations of triplexes. In particular, it was found that respondents residing in central neighbourhoods had almost no aversion to triplexes relative to detached houses (with a coefficient of -0.066 (-1.349 + 1.283)). Those not residing downtown, on the other hand, had an even more negative impression (coefficient of -1.349) than in the text survey. With respect to significance, most of the coefficients are significant at 1% apart from the townhouse coefficient in the gaming survey that is significant at 10%. In addition to being able to segment the triplex coefficient by residential location, there was a significant distribution across the respondents for two dwelling types (triplexes and apartments) in the gaming model, whereas there was no significant distribution for triplexes in the text-only survey.

The results obtained for space between buildings were similar across both surveys and both coefficients were significant at 1%. A significant distribution is observed for this coefficient in both models with the standard deviation of the distribution being a bit larger in the gaming model as well as showing higher statistical significance.

With respect to front yard depth, somewhat different results are found in the two models. Whereas the text-only model found a significant and positive relationship for the entire sample, no significant relationship was observed for the entire sample in the gaming survey. At the same time, a negative and significant (at 5%) coefficient was found for front yard depth among urban core residents. No significant distribution for this coefficient was found in either model.

Coefficients for the three types of travel time were statistically significant and right-sided. The magnitudes were similar between the two models, with travel time to work by car coefficient being almost identical, and the travel time to shops (by foot) smaller in the text-only model. With respect to travel time to work by transit, the coefficient in the text-only model is smaller than the gaming model. Moreover, it was possible to segment travel time to work by transit in the gaming model, which resulted in non-drivers being twice as sensitive to this variable than drivers. It was also possible to estimate significant distributions for travel time to work by car in both models.

With respect to the home value variable, different results were found between the two models. The primary difference is that the coefficient was right-sided and significant for the entire sample in the gaming survey. Moreover it appears that respondents reporting higher base values were less sensitive to home values than those reporting lower (equal or less than $300k) base values. In the text-only model, respondents reporting higher base values did not appear to be sensitive to average home values. Also, average home value was not only significant across the entire sample in the gaming model, but the size of the coefficient was also larger (twice as large for under $300k base price). No significant distribution could be estimated in either of the models.

5.4.2 Discussion of Results

Perhaps the most interesting finding when examining the two models is how similar they actually are. At the same time, after looking at the specifics of model fit and coefficients, it is helpful to examine the results globally to better understand the nuances between the models.

A particularly interesting result comes from the combination of the findings of the individual coefficient results. Previous studies (Jansen et al., 2009; Orzechowski et al., 2005; Vriens et al., 1998) have found that visually presented variables tend to take on more importance than variables presented via text. This happens in two ways. First, variables presented visually have been found to have larger coefficients than when they are presented as text (Jansen et al., 2009; Vriens et al., 1998). Second, variables presented as text in the presence of other visual variables have been found to have lower or nonsensical values (Orzechowski et al., 2005). This does not appear to be the case in our results.

With respect to the coefficients of the visual variables, they are either smaller in absolute value (townhouse and apartment coefficients), or quite close to the value of the coefficient in the text-only model (space between buildings). The one exception to this is the triplex coefficients in the gaming model. The difference here, however, does not appear to suggest that the level "triplex" is taking on greater importance than in the text-only model. Rather, it seems that being able to segment by residential location suggests different preferences for those living in the urban core relative to those that are not. It is worth noting that when estimated without segmentation, the coefficients are practically the same; -0.967 in the text model, and -0.984 in the gaming model. With respect to the coefficients of the text variables in the presence of visual variables, they have similar (travel times to work) or larger values (travel time to shops, average home value) to those in the text-only model. Moreover, the text variables in the gaming model are all statistically significant. This is not the case in the text-only model for average home value for respondents reporting base prices greater than $400k.

There are a couple of possible explanations for our findings compared to what has previously been reported. The first relates to the fact that most of the visual (and in particular the townhouse and apartment) variables have coefficients of smaller absolute values than in the text-only survey. This result amounts to saying that respondents were less averse to these dwelling types than detached houses.

In order to reduce the complexity of the survey, the levels of dwelling type were not modified for any other characteristics. One could imagine, for example, dwelling types of different ages, states of repair and architectural design. The inclusion of such additional attributes would not only cause the design of the 3D environments to be much more onerous, but would also make the experimental design more complex and add additional burden to respondents. Since the aim of the research was to evaluate a gaming platform as a medium by which to administer SP surveys, we wanted to keep the design and number of variables manageable. As such, we decided to choose buildings with specific ages, architectural designs, etc. The designs chosen were contemporary for all of the dwelling types apart from triplexes. Triplexes are the dominant dwelling type of central neighbourhoods in Montreal, and as such also date to the first half of the 20^{th} century. In focus groups it was clear that the survey would appear unrealistic if this type of characteristic dwellings was not included as a level.

We tried to account for the age and condition of dwelling types in the text-only survey by describing them as "recently built" or "recently renovated" (in the case of triplexes). However, given the smaller absolute values of the townhouse and apartment coefficients in the gaming model, it seems that respondent impressions of dwelling types were more negative in the text model than the explicit dwelling types that were presented in the gaming survey. We believe that this helps to explain why the townhouse and apartment coefficients were smaller in absolute value in the gaming model. It also needs to be mentioned that we see this as a disadvantage of using a gaming platform, at least in the context of a design that does not account for additional characteristics of dwelling types – namely that because one has to choose very specific features of the built environment, such as dwelling types, that a certain degree of generalizability of the exercise may be lost. In other words, the specific idea of what dwelling types are like may not be consistent with what people think of dwelling types when they are not prompted visually, and as a result this may lead to different coefficient estimates.

The second possible explanation for the results (i.e. that visual attributes don't seem to take on more importance than the text attributes) is in the design of the gaming survey itself. It seems that the integration of the visual and text information in the choice task options, as well as with the summary screen, has allowed respondents to focus as much on the text as on the visual attributes. This is undoubtedly an advantage.

This last result, combined with the fact that in the gaming model it appears possible to extract more information from respondent data (i.e. additional significant coefficients, the ability to segment by respondent type), and that we have a slightly better model fit, lead us to conclude that the gaming platform was better able to focus respondent attention. This conclusion shouldn't be overstated since the gaming model is not vastly superior to the text-only model, but it does seem to be at least marginally better. As such, while the townhouse and apartment coefficients in the gaming model may underestimate the degree to which respondents are averse to dwelling types other than detached homes, the platform does seem to be better able to focus respondent attention on visual *and* text variables. Perhaps this is a worthwhile trade-off. Moreover, perhaps by better accounting for more dwelling type characteristics, the disadvantages could be overcome.

5.5 CONCLUSION

This paper has discussed the findings of a study employing an innovative presentation methodology for SP surveys in the domain of neighbourhood choice. In particular, it compared a traditional text-only stated-preference survey with a survey administered through a 3D gaming simulation where respondents could navigate through alternative neighbourhoods.

Our main findings are different from what has been found in studies by Jansen et al. (2009) and Orzechowski et al. (2005), in which attributes tended to have greater importance when presented visually. In this study, we did not find that visual attribute coefficients had greater importance than text variables. In the case of the dwelling type (particularly townhouses and apartments), coefficients appear to have been smaller in absolute terms partly because of the actual representation of dwelling types used that may not correspond with respondent ideas of the dwelling types. This is a disadvantage of such an approach. At the same time, text variables do not appear to have reduced importance in the presence of visual attributes, and there is some evidence that the gaming platform was better able to focus respondent attention and process the text as well as visual attributes. In the gaming model it was possible to incorporate more variables, including respondent type segmentation and random parameters for visual attributes thus contributing to an overall better model fit than the text only model.

This research attempted to keep the number of attributes as low as possible in order to produce a manageable SP study. However, further research might focus on determining whether or not having a more complex survey design (e.g. by including other dwelling type variables such as age, condition, etc.) might be able to overcome the disadvantages of such an approach. This would have to be balanced of course with the additional burden this would imply for respondents. Additionally, the inclusion of some attributes that cannot easily be communicated using only words, such as noise for example, could also be tested through the use of this interface. The advances in graphical software have recently made high quality representations relevant for commercial practices. It is surprising that little research has been done in this domain.

5.6 GLOSSARY

Stated preference (SP) survey: survey where respondents are asked to choose between alternatives in hypothetical scenarios (choice tasks) characterized by different attributes.

Choice task: a hypothetical scenario provided by the researcher to be evaluated by respondent in an SP. Choice tasks are generated according to an experimental design and involve two or more alternatives.

Alternative: the different options available to a respondent to choose within a choice task.

Attributes: characteristics of the alternatives within a choice task.

Attribute levels: values of the attributes characterizing alternatives.

Multinomial logit model (MNL): is the simplest and most popular discrete choice model.

Mixed-logit (ML): an extension of the MNL that can overcome three of its limitations by allowing random taste variation, unrestricted patterns, and correlation in unobserved factors.

Gaming engine: a software platform designed for the creation and development of video-games.

5.7 ACKNOWLEDGEMENTS

The authors would like to thank Sydney Swaine-Simon for his hard work in programming the two versions of the survey in Unity. They would also like to acknowledge the contribution of the Canada Research Chairs program and the Canadian Foundation for Innovation to this research.

6. Overall Conclusions and Future Work

Improving stated-preference survey design and the reliability of collected data has been an important stream of research in stated-preference methods. Despite great interest in using visual techniques among both practitioners and scholars, still little is known empirically about their influences on data quality and the advantages and disadvantages of using such techniques in stated-preference surveys. The thesis presented here contributes to this literature by investigating the extent to which the results obtained through an innovative 3D gaming approach and a traditional text-only stated-preference survey of neighbourhood choice differ.

The issues examined in this research include the impact of visualization on the performance of estimated mixed logit models, the amount of error variance in the choice models as well as the differences observed in the level of significance and the values of the estimated coefficients. It is explained that the gaming platform appeared to better focus respondent attention and result in data that led to a better model of neighbourhood choice than a text-only instrument. At the same time, this may be at the expense of losing some degree of generalizability of the conclusions due to providing specific visual information about alternative neighbourhoods.

It is also highlighted in the research that the graphical design of a visual survey is very important. As shown in the work of other researchers (i.e. Jansen et al. (2009); Orzechowski et al. (2005)), respondents might fail to adequately pay attention to either visual or textually presented elements of the survey. Based on my experience with this research, I believe that a proper integration of visual and non-visual attributes is critical when visualization is used in the design of a stated-preference survey. I believe using gaming platforms as a medium could overcome some of these issues by overlaying written information over graphics. Future research though needs to investigate how visual and non-visual information are processed. This could be done by using eye-tracking devices, as seen Boumeester et al. (2008).

Furthermore, a 3D gaming approach, in addition to improving the understanding of respondents by providing visual materials, is capable of conveying hard-to-communicate attributes, such noise or aesthetic parameters. The inclusion of such attributes that cannot easily be communicated using only words, such as noise for example as used in Ortuzar and Rodríguez (2002), could also be tested through the use of this interface. Moreover, testing the applicability of more cutting-edge techniques, such as augmented reality, is another possible direction for future research. This approach, in contrast to virtual reality which completely replaces the real world with a virtual counterpart, may be used to generate more realistic images by adding information to the users view of a real-world scenes (Schall et al., 2011).

While the thesis presented here provides empirical evidence that representation method can affect the nature of results in a stated-preference survey, still, many questions remain unanswered. It is often pointed out in the literature (Dijkstra et al., 1996; Green & Srinivasan, 1978; Vriens et al., 1998; Wittink et al., 1994) that respondents are more likely to make choices as they would in the real-world if survey instruments better mimic real-world choice behaviour. This study, unfortunately, had no access to revealed preference data sets to measure how much the data collected through each survey interface matches with real-world behaviour (i.e. external validity). Comparing the estimated coefficients or values of willingness-to-pay derived from choice models using revealed land use or travel data would have been particularly informative.

Another limitation is that the sample used in this research could not be considered to be truly representative of the population of Montrealers. It would have been interesting, if the sample sizes could have been large enough or of the right composition to be representative of the population of Montreal. With this, it could have been determined whether the estimated models are in close agreement with past stated-preference studies previously performed in Canada.

Another issue in this study is the distributional assumption made with respect to random parameters within the estimated mixed logit models. Distributions are arbitrary approximations to the real behavioural profiles and the researcher selects the one that best matches the behaviour under study. Applying the lognormal distribution to the parameters that needed to have a specific sign (negative as for price and travel time parameters) made model estimation hard in some cases, especially when there were large numbers of Halton draws (e.g. 10,000). The likelihood function for lognormal distribution is extremely flat around the maximum and experience shows that this makes convergence of the function very hard (Algers, Bergström, Dahlberg, & Lindqvist Dillén, 1998; Hensher & Greene, 2003; Sillano & Ortuzar, 2005). Moreover, forcing parameters to follow a lognormal distribution may not be necessary and potential problems with this function could be avoided through using other functions (Sillano & Ortuzar, 2005). As it is proposed in Hensher and Greene (2003), applying truncated or constrained distributions (e.g.

generalized constrained triangular distribution) is an alternative approach to follow in the future analyses of our data sets..

7. References

Algers, S., Bergström, P., Dahlberg, M., & Lindqvist Dillén, J. (1998). Mixed logit estimation of the value of travel time: Working Paper, Department of Economics, Uppsala University.

Arentze, T., Borgers, A., Timmermans, H. J. P., & DelMistro, R. (2003). Transport stated choice responses: effects of task complexity, presentation format and literacy. *Transportation Research Part E: Logistics and Transportation Review, 39*(3), 229-244. doi: 10.1016/s1366-5545(02)00047-9

Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand* (Vol. 9): MIT press.

Bierlaire, M. (2003). *BIOGEME: A free package for the estimation of discrete choice models.* Paper presented at the 3rd Swiss Transportation Research, Ascona, Switzerland.

Boumeester, H., Coolen, H., Dol, C., Goetgeluk, R., Jansen, S., Mariën, A., & Molin, E. (2008). Module Consumentengedrag WoON 2006, Hoofdrapport. *Delft: Onderzoeksinstituut OTB*.

Boyle, K. J., Holmes, T. P., Teisl, M. F., & Roe, B. (2001). A comparison of conjoint analysis response formats. *American Journal of Agricultural Economics, 83*(2), 441-454.

Caussade, S., Ortuzar, J. d. D., Rizzi, L. I., & Hensher, D. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B: Methodological, 39*(7), 621-640. doi: 10.1016/j.trb.2004.07.006

Cooper, J., Ryley, T., & Smyth, A. (2001). Energy trade-offs and market responses in transport and residential land-use patterns: Promoting sustainable development policy. *Urban Studies, 38.9*, 1573-1588. doi: 10.1080/00420980126673

Dijkstra, J., Roelen, W., & Timmermans, H. J. P. (1996). *Conjoint measurement in virtual environments: a framework*. Paper presented at the 3rd Design and Decision Support Systems in Architecture and Urban Planning.

Gan, C. E., & Luzar, E. J. (1993). A conjoint analysis of waterfowl hunting in Louisiana. *Journal of Agricultural and Applied Economics, 25*(02).

Gayda, S. (1998). *Stated preference survey on residential location choice and modal choice in Brussels.* Paper presented at the Transportation Planning Methods. Proceedings of seminar D held at AET european transport conference, Loughborough University, UK, 14-18 September 1998. Volume P423.

Gillespie, J., Taylor, G., Schupp, A., & Wirth, F. (1998). Opinions of professional buyers toward a new, alternative red meat: Ostrich. *Agribusiness, 14*(3), 247-256.

Green, P. E., & Rao, V. R. (1971). Conjoint Measurement for Quantifying Judgmental Data. *Journal of Marketing Research, 8*(3), 355-363. doi: 10.2307/3149575

Green, P. E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: issues and outlook. *consumer research*, 103-123.

Green, P. E., & Srinivasan, V. (1990). Conjoint analysis in marketing: new developments with implications for research and practice. *The Journal of Marketing*, 3-19.

Harrison, R. W., Stringer, T., & Prinyawiwatkul, W. (2002). An analysis of consumer preferences for value-added seafood products derived from crawfish. *Agricultural and Resource Economics Review, 31*(2), 157-170.

Hensher, D. (1994). Stated preference analysis of travel choices: the state of practice. *Transportation, 21*(2), 107-133. doi: 10.1007/bf01098788

Hensher, D., & Greene, W. (2003). The mixed logit model: the state of practice. *Transportation, 30*(2), 133-176.

Hensher, D., Louviere, J., & Swait, J. (1998). Combining sources of preference data. Journal of Econometrics. *Econometrics, 89*(1-2), 197-221.

Hensher, D., Rose, J., & Greene, W. (2005). The implications on willingness to pay of respondents ignoring specific attributes. *Transportation, 32*(3), 203–222.

Holbrook, M., & Moore, W. (1981). Feature Interactions in Consumer Judgments of Verbal versus Pictorial Presentations. Journal of Consumer Research. *Consumer Research, 8*(1), 103-113.

Holland, D., & Wessells, C. R. (1998). Predicting consumer preferences for fresh salmon: the influence of safety inspection and production method attributes. *Agricultural and Resource Economics Review, 27*, 1-14.

Hunt, J. D. (2001). Stated Preference Analysis of Sensitivities to Elements of Transportation and Urban Form. *Transportation Research Record, 1780*(1), 76-86.

Hunt, J. D., McMillan, J. D. P., & Abraham, J. E. (1996). Stated preference investigation of influences on attractiveness of residential locations. *Transportation Research Record, 1466*, 79-87.

Jaeger, S. R., Duncan, H., & MacFie, H. J. H. (2001). Methodological issues in conjoint analysis: A case study. *European Journal of Marketing, 35*(11/12), 1217-1237.

Jansen, S., Boumeester, H., Coolen, H., Goetgeluk, R., & Molin, E. (2009). The impact of including images in a conjoint measurement task: evidence from two small-scale studies. *Housing and the Built Environment, 24*(3), 271-297. doi: 10.1007/s10901-009-9149-x

Kim, J. H. (2006). Amenity Valuing Differentiation in Residential Location Choice among Income Groups: A Stated Preference Approach. *International Journal of Urban Sciences, 10*(1), 41-57. doi: 10.1080/12265934.2006.9693586

Kim, J. H., Pagliara, F., & Preston, J. (2005). The intention to move and residential location choice behaviour. *Urban Studies, 42*(9), 1621-1636. doi: 10.1080/00420980500185611

Levine, J., & Frank, L. D. (2006). Transportation and land-use preferences and residents' neighborhood choices: the sufficiency of compact development in the Atlanta region. *Transportation, 34*(2), 255-274. doi: 10.1007/s11116-006-9104-6

Lin, B.-H., Payson, S., & Wertz, J. (1996). Opinions of professional buyers toward organic produce: A case study of mid-Atlantic market for fresh tomatoes. *Agribusiness, 12*(1), 89-97.

Louviere, J. J., Flynn, T. N., & Carson, R. T. (2010). Discrete Choice Experiments Are Not Conjoint Analysis. *Journal of Choice Modelling, 3*(3), 57-72. doi: http://dx.doi.org/10.1016/S1755-5345(13)70014-9

Louviere, J. J., Hensher, D., & Swait, J. (2000). *Stated choice methods: analysis and applications*: Cambridge University Press.

Louviere, J. J., & Timmermans, H. J. P. (1990a). Hierarchical Information Integration Applied to Residential Choice Behavior. *Geographical Analysis, 22*(2), 127-144.

Louviere, J. J., & Timmermans, H. J. P. (1990b). Stated preference and choice models applied to recreation research: a review. *Leisure Sciences, 12*(1), 9-32.

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology, 1*(1), 1-27. doi: http://dx.doi.org/10.1016/0022-2496(64)90015-X

Miguel, F. S., Ryan, M., & McIntosh, E. (2000). Applying conjoint analysis in economic evaluations: an application to menorrhagia. *Applied Economics, 32*(7), 823-833.

Molin, E. (2011). Conjoint analysis *The measurement and analysis of housing preference and choice* (pp. 127-155): Springer.

Molin, E., & Timmermans, H. J. P. (2003). *Accessibility Considerations in Residential Choice Decisions:Accumulated Evidence from the Benelux*. Paper presented at the 82nd Annual Meeting of the Transportation Research Board Washington, DC.

Morrow-Jones, H., Irwin, E., & Roe, B. (2004). Consumer Preference for Neotraditional Neighborhood Characteristics. *Housing Policy Debate, 15*(1), 171-202. doi: 10.1080/10511482.2004.9521498

MVA Consultancy, ITS University of Leeds, & TSU University of Oxford. (1987). Value of Travel Time Savings. *Policy Journals, Newbury, Berks*.

Ortuzar, J. d. D., Martinez, F. J., & Varela, F. J. (2000). Stated Preferences in Modelling Accessibility. *International Planning Studies, 5*(1), 65-85. doi: 10.1080/135634700111828

Ortuzar, J. d. D., & Rodríguez, G. (2002). Valuing reductions in environmental pollution in a residential location context. *Transportation Research Part D: Transport and Environment, 7*(6), 407-427. doi: http://dx.doi.org/10.1016/S1361-9209(02)00010-X

Ortuzar, J. d. D., & Willumsen, L. G. (1994). *Modelling transport* (fourth ed.): John Wiley & Sons, Ltd.

Orzechowski, M. A., Arentze, T. A., Borgers, A. W. J., & Timmermans, H. J. P. (2005). Alternate methods of conjoint analysis for estimating housing preference functions: Effects of presentation style. *Housing and the Built Environment, 20*(4), 349-362. doi: 10.1007/s10901-005-9019-0

Prentice, B. E., & Benell, D. (1992). Determinants of empty returns by US refrigerated trucks: conjoint analysis approach. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 40*(1), 109-127.

Ramírez-Hurtado, J. M. (2010). Measuring preferences: From conjoint analysis to integrated conjoint experiments. *Revista de Métodos Cuantitativos para la Economía y la Empresa, 9*, 28-43.

Rid, W., & Profeta, A. (2011). Stated Preferences for Sustainable Housing Development in Germany-A Latent Class Analysis. *Journal of Planning Education and Research, 31*(1), 26-46. doi: 10.1177/0739456x10393952

Rizzi, L. I., Limonado, J. P., & Steimetz, S. S. C. (2012). The impact of traffic images on travel time valuation in stated-preference choice experiments. *Transportmetrica, 8*(6), 427-442. doi: 10.1080/18128602.2010.551524

Sawtooth-Software-Inc. (2013). Sawtooth Software: The CBC System for Choice-Based Conjoint Analysis (Version 8).

Schall, G., Schöning, J., Paelke, V., & Gartner, G. (2011). A survey on augmented maps and environments: Approaches, interactions and applications. *Advances in web-based GIS, mapping services and applications. Taylor & Francis Group, UK*.

Senior, M. L., Webster, C. J., & Blank, N. E. (2006). Residential relocation and sustainable urban form: Statistical analyses of owner-occupiers' preferences. *International Planning Studies, 11*(1), 41-57. doi: 10.1080/13563470600935024

Sillano, M., & Ortuzar, J. d. D. (2005). Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning A, 37*(3), 525-550.

Statistics-Canada. (2003). CHASS data centre.

Swait, J., & Louviere, J. J. (1993). The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Marketing Research, 30*(3), 305-314.

Train, K. E. (2009). *Discrete choice methods with simulation*: Cambridge university press.

Vriens, M., Loosschilder, G. H., Rosbergen, E., & Wittink, D. R. (1998). Verbal versus Realistic Pictorial Representations in Conjoint Analysis with Design Attributes. *Product Innovation Management, 15*(5), 455-467.

Walker, B., Marsh, A., Wardman, M., & Niner, P. (2002). Modelling Tenants' Choices in the Public Rented Sector: A Stated Preference Approach. *Urban Studies, 39*(4), 665-688. doi: 10.1080/00420980220119516

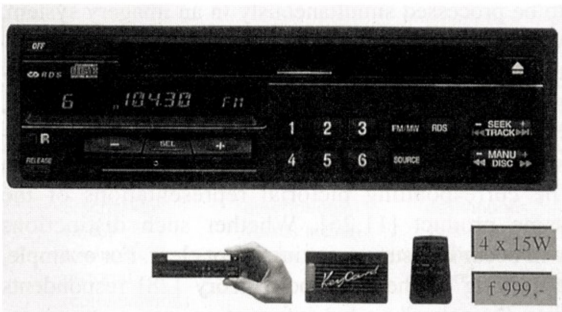
Wang, D., & Li, S.-M. (2004). Housing preferences in a transitional housing system: the case of Beijing, China. *Environment and Planning A, 36*(1), 69-87.

Wittink, D. R., & Cattin, P. (1989). Commercial Use of Conjoint Analysis: An Update. *Journal of Marketing, 53*(3), 91-96. doi: 10.2307/1251345

Wittink, D. R., Vriens, M., & Burhenne, W. (1994). Commercial use of conjoint analysis in Europe: Results and critical reflections. *International journal of Research in Marketing, 11*, 41-52.

8. Appendices

Appendix 1 Example of a choice scenario with traffic images in Rizzi et al. (2012)



A

Profile 1

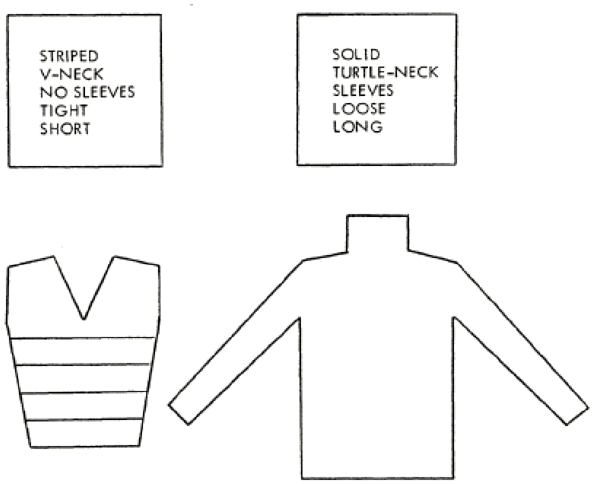
Amplifier control mode	Rotary control
Remote control	Present
CD changer	Present
Indash player system	Tape player with Dolby B noise reduction
Amplifier power	4×20 Watt
Anti-theft system	Detechable front
Price	NLG 1399.=

B

Profile 2

Amplifier control mode	Touch control
Remote	Present
CD changer	Absent
Indash player system	CD player
Amplifier power	4×15 Watt
Anti-theft system	Detachable front and key card
Price	NLG 999.=

Appendix 2 Examples of written versus pictorial representations in Vriens et al. (1998)



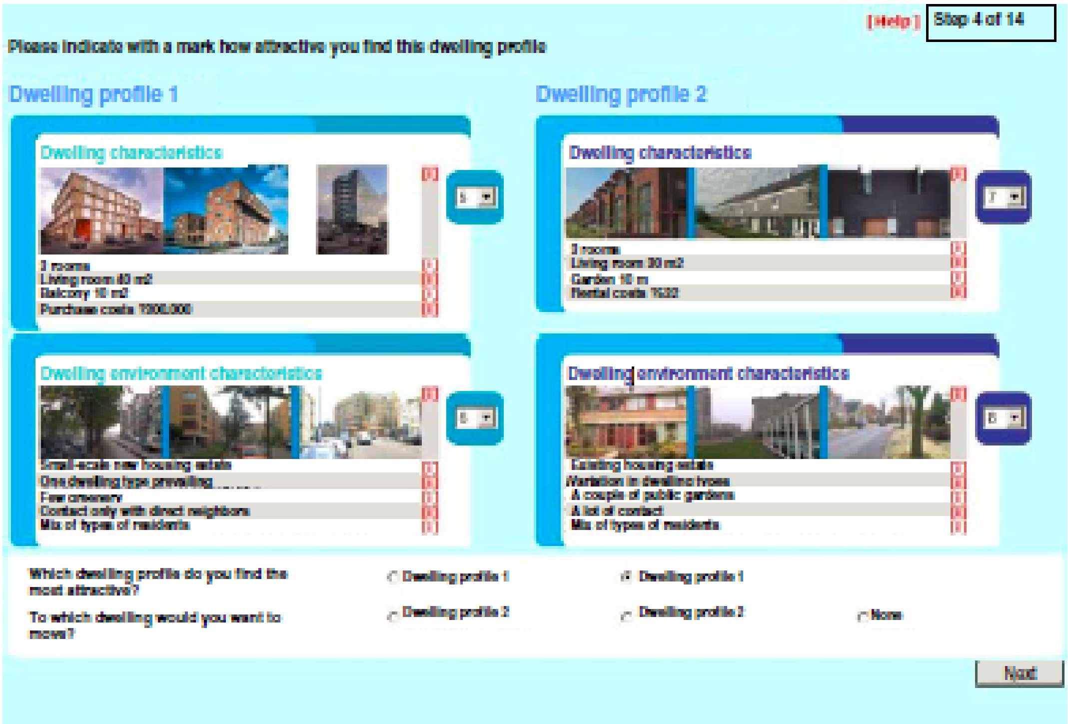
Appendix 3 Examples of Verbal descriptions versus pictorial display used in Holbrook and Moore (1981)

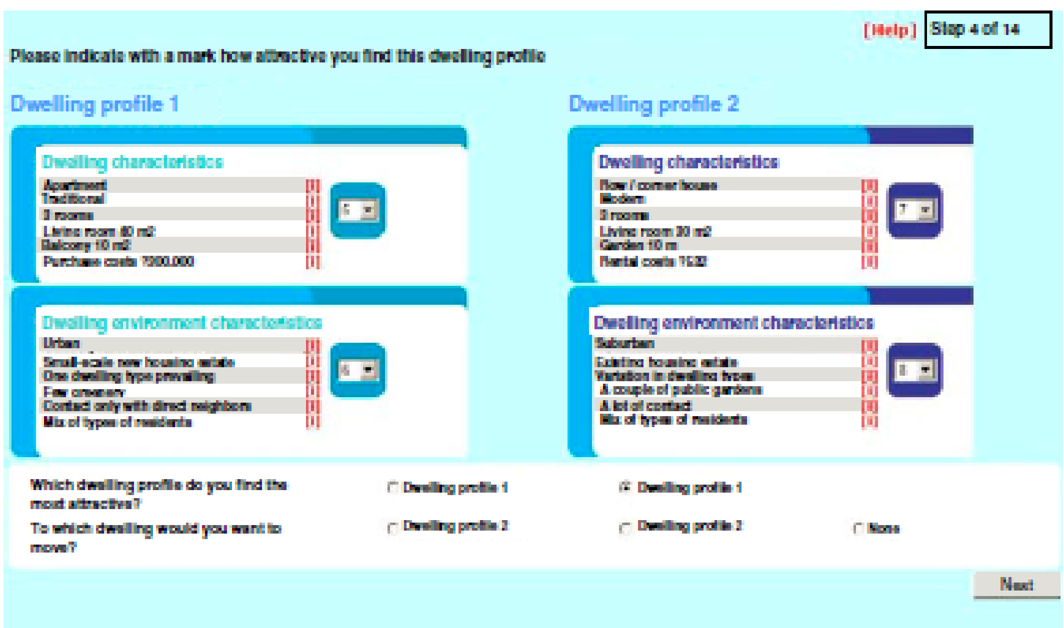
Type of dwelling	Semi-detached dwelling
Style	Modern architectural style
Environment	Urban residential environment
Monthly costs	€ 500
Rooms	5 rooms

Type of dwelling	Semi-detached dwelling
Style	Modern architectural style
Environment	Urban residential environment
Monthly costs	€ 500
Rooms	5 rooms

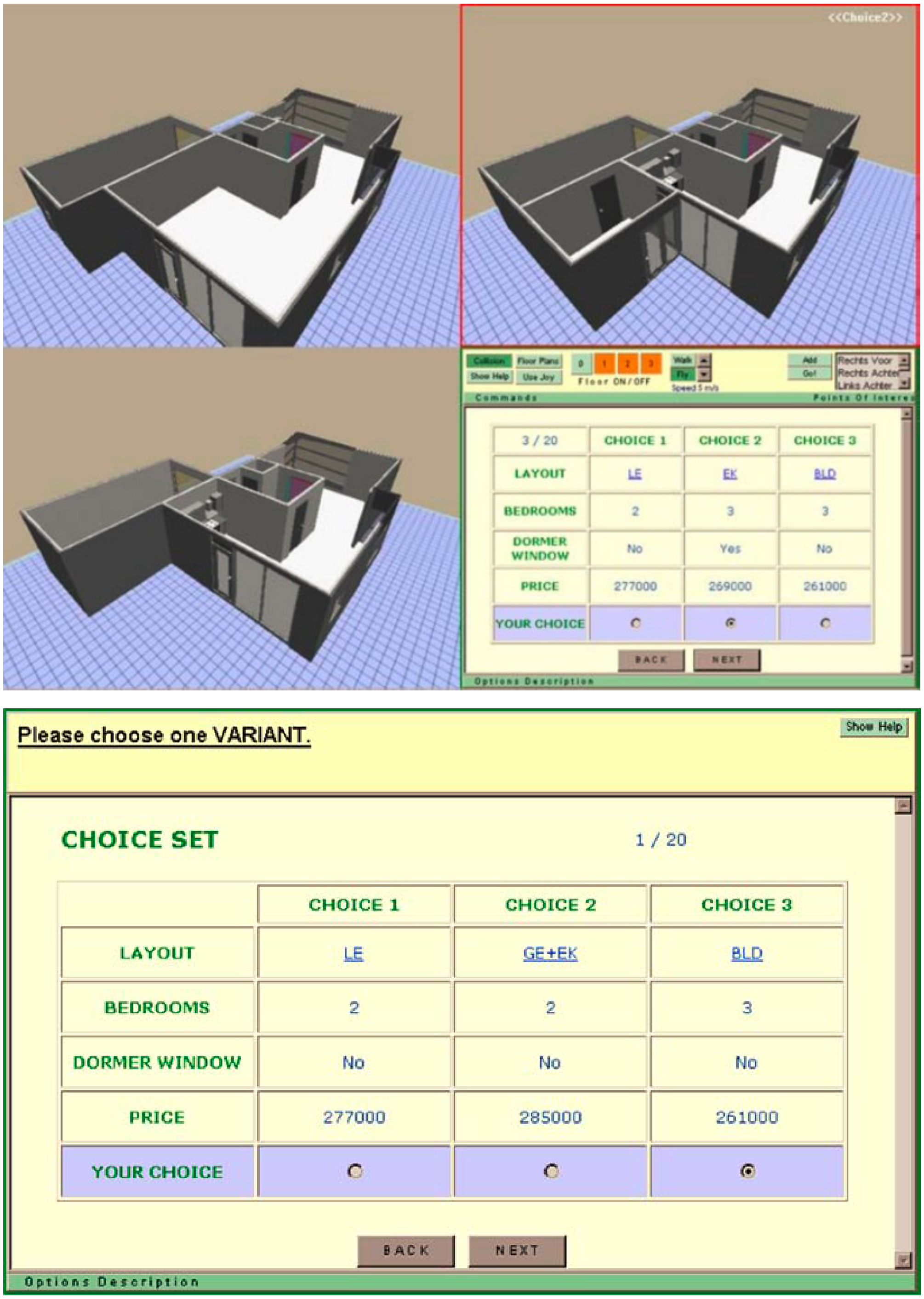
Type of dwelling	Semi-detached dwelling
Style	Modern architectural style
Environment	Urban residential environment
Monthly costs	€ 500
Rooms	5 rooms

Appendix 4 Examples of the instruments used in study 1 of Jansen et al. (2009)





Appendix 5 Example of the instruments used in study 2 of Jansen et al. (2009)



Appendix 6 Multimedia instrument (above) versus the text-only instrument (below) used in Orzechowski et al. (2005)

Local air quality:		noticeably bad 1 day per month
Municipal taxes or rent:		**up $125 per month (up $1500 per year)**
Traffic noise:		occasionally just noticeable
Street in front of dwelling:		**local road**
Trip to work:	By car:	**20** minutes in vehicle **$2.00** per day for parking **$0.75** for fuel & user charges (fuel taxes & road tolls)
	By transit:	45 minutes in vehicle with no transfers 3 block walk to bus stop $0.75 fare one way
Trip to shop:	By car:	15 minutes in vehicle $0.50 for parking $2.00 for fuel & user charges (fuel taxes & road tolls)
	By transit:	30 minutes in vehicle with no transfers 1 block walk to bus stop $0.75 fare one way
Walk to local elem school:		**5 minutes** **using crosswalk on collector road with LRT**
Dwelling type:		**single family**

Appendix 7 An example of a hypothetical stated-preference card in Hunt (2001)

If I were to move, I'd like to find a neighborhood . . .

A. with larger homes, even if this means I have to drive for all of my trips.	*or*	**B.** where I can walk, bicycle or take public transit for some of my trips, even if this means that the homes are smaller.

17) *Your neighborhood preference is:*

0	1	2	3	4	5	6	7	8	9	10
strongly prefer **A**			somewhat prefer **A**		neutral		somewhat prefer **B**			strongly prefer **B**

---

18) *Please indicate whether your current neighborhood is more like "A" or "B":*

0	1	2	3	4	5	6	7	8	9	10
more like **A**					equally like A & B					more like **B**

---

19) *Regarding house size and travel options, the neighborhood you'd hope to find would be:*

0	1	2	3	4	5	6	7	8	9	10
more like **A** than your current neighborhood					like your current neighborhood					more like **B** than your current neighborhood

Appendix 8 An example of Levine and Frank (2006) stated-preference survey