How Do Visual Representations Influence Survey Responses?

Evidence from a Choice Experiment on Landscape Attributes of Green Infrastructure

Yau-Huo (Jimmy) Shr

Center for Agricultural and Rural Development, Iowa State University

Richard Ready

Department of Agricultural Economics and Economics, Montana State University

Brian Orland

College of Environment+Design, University of Georgia

Stuart Echols

Department of Landscape Architecture, Pennsylvania State University

Manuscript Correspondence:

Yau-Huo Shr

yhshr@iastate.edu

814-777-0087

Abstract

This article provides new evidence on how images influence survey responses, using a split-sample choice experiment. Our results suggest that, when respondents are presented with both images and text, they exhibit stronger preferences for attributes with high visual salience than when presented with either images or text alone. Furthermore, respondents are less likely to ignore individual attributes when both images and text are provided. However, the provision of images makes responses more random, i.e., respondents' preferences for attributes are less consistent across choice questions.

Keywords: Non-market valuation, Generalized mixed logit model, Cognitive overload, Landscape visualizations

---

The authors thank Abhinandan Bera, Department of Landscape Architecture, Pennsylvania State University, for generating the computer-rendered images used in this study. This study was financially supported by the US EPA Star Grant: Center for Green Infrastructure and Stormwater Management (Funding Number: EPAG2011-STAR-A1).

Introduction

Visual representations such as photographs and computer-generated images have been widely used to illustrate choice questions and ensure respondents understand the scenarios in stated preference studies. Using visual representations has been recognized as one of the most effective approaches to promote the comprehension or evaluability of the studied objects (Bateman et al., 2009; Mathews et al., 2006). In spite of their widespread use, there are only a small number of studies that have explored whether the use of visual representations affects the results of stated preference methods compared to using other methods (e.g., text descriptions) of communicating the attributes that respondents are being asked to value. We found four studies in the field of environmental economics that have examined the use of novel dynamic visual representations, including 3D films or virtual environment on results of choice experiments (Bateman et al., 2009; Matthews et al., 2017; Patterson et al., 2017; Rid et al., 2018), but how widely-used static visual representations affect people's responses is largely unexplored.

This paper uses a split-sample choice experiment with three alternative survey treatments to provide evidence on how people's choices are influenced by visual representations and focuses on the impacts of using static, computer-rendered images. The choice experiment in this paper values landscape attributes of green infrastructure (GI), which has been promoted to combat severe environmental problems such as downstream water pollution and urban flooding caused by stormwater runoff (USEPA, 2016).[1] The three survey treatments provide different information to present choice scenarios – the baseline treatment uses both static images and text descriptions of the options, the text-only treatment uses only text, and the image-only treatment uses solely the images. The impacts

[1] Green infrastructure is a cost-effective stormwater management approach that focuses on creating local-scale ecosystems to treat stormwater at the source and serves as a substitute for or a complement to conventional stormwater management approaches. It generates various benefits for climate resiliency, habitat and wildlife, and human communities (USEPA, 2016).

of showing images, text, or both can therefore be identified through paired comparisons between the baseline and two other treatments.

Among the studies that investigate how 3D films or virtual environments influence respondents' choices, Bateman et al. (2009) and Matthews et al. (2017) both find that respondents made more consistent (less random) choices with the aid of dynamic visual representations. However, Rid et al. (2018) find that choices are less consistent when 3D film sequences were provided, compared to using static images. Patterson et al. (2017) suggest that the use of a virtual environment results in better respondent attention and more significant marginal utility coefficients.[2]

In light of the findings from these existing studies on the use of dynamic images, we propose the following three hypotheses regarding the use of static images: providing static images, in addition to text descriptions, will (1) change people's taste, i.e., the marginal utility or willingness to pay estimates, (2) better focus respondents' attention on the attributes, and (3) reduce randomness in making choices. We apply generalized mixed logit models (GMXL) (Fiebig et al., 2010),[3] which allow us to test these hypotheses simultaneously. GMXL extends one of the current standard discrete choice models, the mixed (or random parameter) logit model, by accommodating the differences in the randomness of people's choices (i.e., scale heterogeneity). This model enables the measurement of choice randomness via the scale parameter in random utility model (the variance of the error term in utility function).

Our paper contributes to the literature by investigating the impacts of using static images on respondents' choices and providing caveats for scholars when conducting stated preference research that includes static visualizations. We find evidence supporting the first

[2] Among transportation literature, Harline and Burris (2014) and Rizzi et al. (2012) both study the impacts of traffic photos on the value of time.

[3] The model is called the generalized multinomial logit model in the terminology of Fiebig et al. (2010), but we prefer the name GMXL suggested by Greene and Hensher (2010) because GMXL is a general form of the mixed logit.

two hypotheses but the contrary of the third hypothesis. The findings indicate that the policy implications of research can be greatly influenced by the method of describing the survey questions.

A second contribution of our paper is that it provides information on preferences over design aspects of GI. When it comes to the adoption of new technologies that generate environmental benefits such as solar panels, wind farms, and GI, the appearance of such infrastructure is one of the principal determinants of public acceptance (del Carmen Torres-Sibille et al., 2009; Dimitropoulos and Kontoleon, 2009). The environmental benefits of decentralized stormwater management have been examined together with flood frequency reductions by Londoño Cadavid and Ando (2013). However, the understanding of what types of GI landscape design are preferred by citizens remains sparse (Tzoulas et al., 2007; Schäffler and Swilling, 2013).

This article proceeds with a review of the use of visual representations and the potential impacts on people's choices. An overview of the survey design and data is presented before the econometric treatment. The empirical results section first provides the overall findings regarding preferences for the landscape attributes of GI, and then examines the impacts of visualization on people's tastes, choice randomness, and attention to the questions. The final section summarizes the findings and provides suggestions for future research.

Visual Representations and Choices

Visual representations may be better at capturing resource conditions that would be difficult to describe using narrative methods (Manning and Freimund, 2004). Photo-based surveys have demonstrated their suitability as an instrument when evaluating landscape preferences (Barroso et al. 2012), and preferences solicited using photo-based off-site surveys are found to be highly consistent with those from on-site surveys (Hill and Daniel, 2007; Natori and

Chenoweth, 2008). The influence of how and how much information is provided on people's responses to questions regarding their landscape preferences has been a popular research topic outside the discipline of economics, such as in environmental psychology and landscape planning. For example, both Hill and Daniel (2007) and van der Wal et al. (2014) found that respondents' expressed landscape preferences were malleable depending on the additional information they were provided.[4]

Showing images in addition to text descriptions enriches the information set, potentially enabling respondents to make more precise decisions with greater available information. Among stated preference studies, Haider et al. (1998) and Haider and Hunt (2002), in their studies on the visual aesthetic of forested shorelines, are some of the first to use digitally calibrated images. Some recent choice experiment research using static visualizations include Beharry-Borg and Scarpa (2010) on coastal water quality change, Berninger et al. (2010) and Elsasser et al. (2010) on forest conversion programs, De Valck et al. (2014) on shared trail design, Iglesias et al. (2013) on neighborhood design for crime prevention, Ryffel et al. (2014) on land use management on flood protection, and Scarpa et al. (2007) on landscape improvement schemes. However, as we noted earlier, no study has been devoted to investigate the impacts of using static visualizations.

Although visual representation is preferable to text descriptions in many circumstances, visual representations could make consumers use less systematic cognitive processes when making choices (Townsend and Kahn, 2014). A choice task with more sources of information is also a more complex choice task. More complicated choice questions are harder to answer, so people may make their decisions more randomly or exhibit less certainty in picking the option they really prefer. Information overload has long been recognized as a challenge that

[4] These studies use the term "preference" to infer concepts such as judgments on landscape scenic beauty and responses to questions soliciting their preferred landscape conditions, so people's "preferences" are certainly malleable in this context.

can lead to dysfunctional preferences and inconsistent choices (for a recent review, see Chernev et al., 2015). Therefore, although we hypothesize providing static visualizations will reduce the randomness in making choices, in large part because of the findings from studies with similar settings, we consider it is also very plausible to find the opposite. In light of this line of research,[5] we explicitly model the impacts of using supplemental images on both people's taste and randomness in answering choice questions.

Study Design and Data Description

The data used in this study is from a survey of citizen preferences for landscape attributes of green infrastructure. The survey provided respondents with background information about GI and described four core landscape design elements in GI – diversity in plant species, presence of water, percentage of green space mowed, and pattern of plantings (natural vs. designed appearance of plantings).

Respondents were asked to imagine they had decided to move to a new home and were choosing a neighborhood in which to live. Respondents were first asked to choose the housing density of the neighborhood they would likely move into.[6] In subsequent choice questions, visual representations showed a density of built structures that matched the density chosen by the respondent. Twelve choice questions were then presented in randomized order to reduce the potential biases from ordering effect and respondent fatigue (Campbell et al., 2015). Each choice question asked the respondents to choose between a

[5] Although not focusing on the impacts of using visual representations, two recent papers using stated choice experiments have found significant effects on choice certainty when respondents are provided with different information (Czajkowski et al., 2016; Lusk et al., 2008). Both papers adopted the treatment of providing more complete and generally more positively framed information than offered in their control, but their conclusions regarding the impact on respondent choice certainty are mixed.

[6] We provide three housing densities for respondents: detached houses on medium-size lots, detached houses on small-size lots, and townhouses and duplexes.

pair of neighborhoods that have various landscape attributes associated with GI and cost.[7] The choice questions were followed by several debriefing questions regarding respondents' attitudes toward green spaces, their potential concerns, and general expectations.

The levels of four landscape attributes, which were identified through discussion with landscape architects, are listed in Table 1.[8] Diversity in plant species is used to the represent the commonly discussed concept of biodiversity, because it is a dimension of biodiversity that can be nearly perfectly controlled by landscape design and is easy for respondents to understand. Still, the survey does mention that green spaces with greater plant variety will tend to have more varied types of birds, insects, and other animals. The presence of water describes how long water will stay at the designed area after a rainstorm. "Always" indicates that there is always a pool of water (in a retention basin) in the area, though the depth of the water varies over time. The level of "sometimes" is defined as water staying in the retention area up to one day after the storm. "Never" means that water will run off immediately. The percentage of green space mowed is the percentage of *public* grass-covered area that is mowed. The pattern of plantings describes whether the edges of plantings are clearly defined and is a surrogate measure for the level of visible human intervention in shaping the landscape outcomes. "Formal" means that the plantings in a green space are placed in a neat pattern with well-defined edges. "Intermediate" shows that the edges are somewhat defined, while "informal" indicates that there are no apparent edges in planting, so the landscape has

---

[7] We did not offer the option of opt-out or choosing status quo, because this study was not designed to assess the willingness to pay for moving from respondents' current place of living to another neighborhood with some new design of green infrastructure. For that reason, we are only able to measure willingness to pay for marginal changes in GI design. Indeed, from a practical perspective, it is not clear what an opt-out option would look like in our context.

[8] An efficiency design with non-informative prior was used to create 12 choice sets using the SAS %*ChoicEff* macro. A two-block design with 24 choice sets was also tested under random parameter logit specifications, and the simulation result did not show that this 24-choice design was superior to the 12-choice design. Such a design could be inadequate for identifying complex models with large number of parameters. However, we ran multiple models with sparser specifications, such a multinomial logit, and a random parameter logit with/without correlation between parameters, and the results are not qualitatively different.

a naturalistic appearance.[9] A cost attribute is included in the form of an additional annual homeowners' association fee.

We used Visual Nature Studio (VNS), a landscape visualization software, to generate the images describing the neighborhoods in each choice scenario. The terrain of the neighborhood was the same in all visualizations, and was based on the terrain of an actual greenspace used for stormwater control. Parameters were specified that determined the levels of each landscape attribute. Three images from different neighborhood view angles, each from the perspective of a standing person of average height, were created for each option to give respondents an idea of the place as a whole. An example choice question for the base survey version is shown in Figures 1A, 1B, and 1C for the three different neighborhood densities.[10] It is worth pointing out that different techniques of visualization (e.g., photos vs. computer-rendered images) can have various impacts on how respondents answer the questions. Although computer-rendered images do not look as "real" as actual pictures or photographs and may suffer from biased influence of a professional designer, they allow more flexibility in systematically varying attributes while holding terrain constant.

It is a challenge to show intermittent water in static images.[11] In our discussions with stormwater engineers, we learned that features designed to hold water intermittently will look different from features designed to hold water permanently, such as the edge of water or plants in the retention area. We attempted to make the differences in visual representations realistic, although they do not in and of themselves suggest the temporal

---

9 It is worth mentioning that no matter green spaces look highly "man-made" or display natural forms, they are always "designed" by a human.

10 Within each housing density, we assigned the same parameters to VNS. The program assigned the specific locations of some landscape patterns. For this reason, there are some random differences in the patterns of some visual cues (e.g., flowers and trees).

11 The representation of temporal change is a classic challenge in using visualization of dynamic effects, while the use of surrogate static imagery for such situations is well-established. A simple response to the challenge might be to suggest video imagery, but then the issues of time compression and discounting emerge, neither of which has been adequately resolved in the literature. (Hull and Revell, 1989; Williams et al., 2007).

aspect. Another issue is that the differences in the percentage of area mowed could only be easily identified visually from a few specific locations or angles (e.g., from a high elevation or standing right on the mowed area). However, those angles were not able to illustrate the entire landscape. A second complication is that the percentage mowed attribute applies only to the commonly-owned greenspace, while the pictures include views of some privately-owned, mowed land. While this does more closely correspond to conditions found in a natural setting, a change in the percent of commonly-owned land results in a smaller change in the total area mowed (privately-owned and commonly-owned together). The three angles chosen, which were used for all visual representations, do a good job representing the landscape as a whole, although they do not clearly show differences in the percentage of publicly-owned land that is mowed.

This study develops three alternative survey treatments to test our research hypothesis. Hereafter, the three treatment versions are called base, text-only, and image-only. The base version presents the choice scenarios with both images and text descriptions of the attributes, with all images using the housing density chosen by the respondent.[12] The image-only and text-only versions are designed identically to the base version, except for the presentation of the choice scenarios. The text-only version only includes written depictions of each options attributes, and the image-only version only displays only images to illustrate the landscape attributes of the options, with the exception of the cost attribute that is textually stated. Note that, our experiment is not designed to examine how people's responses deviate from their real or "benchmark" behavior with different information treatments, although one can conjecture that the base version treatment is most similar to an actual decision context when choosing a new neighborhood to live in.

Participants were recruited through the KnowledgePanel, a web-panel service provided

[12] There may be differences in preferences among people who choose different housing densities, but that heterogeneity is beyond the scope of this study.

by GfK Knowledge Networks.[13] In total, 499 valid completed surveys were collected in March of 2016, with 159, 170, and 170 responses for the base, text-only, and image-only version, respectively.[14] The participants are residents who live in the suburban areas of the Chesapeake and Delaware Watersheds and are between 25 and 64 years old. We targeted this population because they are more likely to face the context of our choice experiment questions: imagining they had decided to move to a new home and were choosing where to live. The survey regions were filtered by population density at the zip code level to ensure that the survey setting was familiar and understandable for respondents. Only people who lived in a zip code with a population density of 500 – 5000 per square mile were targeted for recruitment. Detailed socio-demographic information of respondents, such as income, education, employment status, home ownership, and marital status, were provided through the KnowledgePanel database. A pretest with 131 panelists who lived in the same survey areas was conducted to identify the survey's potential flaws and ensure ultimate data quality. The pretest responses were not included in the analysis.

Table 2 provides summary statistics for socio-demographic variables and the proportion of respondents choosing the low density neighborhood for all three survey treatments. The weights were provided by the KnowledgePanel and were included in the estimation.[15] Chi-square tests fail to reject the null hypothesis that the distributions of the

---

[13] The households in KnowledgePanel were randomly chosen and the number of surveys in which they are allowed to participate is limited. In addition, KnowledgePanel provides computer and internet services for households without home internet access. These features allow KnowledgePanel to cover more than 95% of US households, and the sample representativeness is thus comparable to that using random digit dialing with cellphone samples.

[14] After completing the fieldwork, 530 people completed the survey with a median completion time of 11 minutes. The samples were trimmed if a respondent chose the same option across all choice questions or completed the survey in less than five minutes. Based on timed trials, we consider it unlikely that a respondent who completed the survey in five minutes indeed answered the questions seriously and carefully.

[15] The socio-demographic variables used for generating the weights across the entire KnowledgePanel include gender, age, education, household income, and home ownership status. For this study, the KnowledgePanel also created the pre-weights of the assigned samples. The pre-weights were designed to make total respondents match the 25-64 year old benchmarks in our designated zip codes for the entire geography of the study. Detail weighting strategy are available in the field report upon request.

variables are the same across the survey version samples, indicating that the randomization was successful and no difference in preferences should be attributed to the differences between the socio-demographic profiles and the chosen housing density of the samples in each version.

Econometric Model

The conceptual framework of stated choice experiments is based on the random utility maximization model (Holmes and Adamowicz, 2003). Preference heterogeneity across individuals has been commonly modeled in terms of people's tastes and randomness when making choices, where the former focuses on the heterogeneity in the marginal utility and the latter is reflected in the scale parameter. In essence, the scale parameter measures the variance in the error term for the random utility function (Hensher et al., 1998; Louviere et al., 1999). The higher (lower) the scale parameter in a respondent's utility function, the lower (higher) the variance in utility over choice options, so a respondent makes less (more) random choices from the perspective of the random utility model.

Formally, a utility (U) function of person i for choosing a certain option q in a choice scenario n, with both the taste and scale heterogeneity built in, can be written as

$$U_{iqn} = [\sigma_{in}\beta + \gamma\Gamma\nu_i + (1-\gamma)\sigma_{in}\Gamma\nu_i]A_{qn} + \varepsilon_{iqn} \tag{1}$$

This is the utility in a generalized mixed logit model (Fiebig et al., 2010). A_{qn} is a vector of the attributes of option q in scenario n, β is a vector of the taste parameters (i.e., the marginal utility of each attribute), and ε_{iqn} is an unobserved random component assumed to be independently and identically distributed following type I extreme value distribution. σ_{in} is the individual- and scenario-specific scale parameter of person i, Γ stands for the lower triangle of the Cholesky matrix, ν_i is a vector of random variables with zero means and known variances representing the individual deviations from the population means of the taste parameters , and γ controls how the covariance matrix of the taste parameters is scaled.[16] We can test our first hypothesis: providing static images, in addition to text descriptions, will change people's taste, by examining treatment-specific βs.

Because σ_{in} and β always enter the utility function multiplicatively, a convenient normalization method of achieving the identification requirement is to assume that σ_{in} follows a lognormal distribution with mean equal to one and variance equal to τ (Fiebig et al., 2010). That is

$$\sigma_{in} = \exp(\bar{\sigma} + \tau\varepsilon_{0in}) \tag{2}$$

where $\bar{\sigma} = -\tau^2/2$ for $\mathrm{E}(\sigma_{in}) = 1$, and $\varepsilon_{0in} \sim N(0, 1)$.

The concept of investigating the differences in scale across individuals, choice scenarios, or survey treatments using the GMXL model was set out in Fiebig et al. (2010). Czajkowski et al. (2016) implement the concept and extend the model by adding the ability to capture observed differences in the scale variance, τ. Following their formulation allows the scale parameter in equation 2 to be rewritten as

$$\sigma_{in} = \exp(\bar{\sigma} + \theta z_{in} + \exp(\lambda z_{in})\tau\varepsilon_{0in}) \tag{3}$$

where z_{in} is a vector of the covariates for the observed heterogeneity in scale mean and variance. In order to control for the observed changes in scale associated with the provision of different information across survey treatments, the focus of z_{in} in our study is as a set of dummy variables indicating which survey treatment a respondent takes. Therefore, a positive (negative) θ shows that the average scale in the utility of samples in the alternative treatment is higher (lower) than that in the control, which indicates that respondents give

[16] The case where γ equals to one, which indicates that the covariance matrix is unaffected by the scale, is the Type I GMXL in the terminology of Fiebig et al. (2010). Alternatively, the Type II GMXL is when γ equals zero, in which the covariances will be as equally scaled as the means. Empirically, γ can be estimated via a logistic transformation as between 0 and 1 or constrained according to the study objectives. For example, Greene and Hensher (2010), Hensher et al. (2011), and Czajkowski et al. (2015) all employed Type II GMXL, although none of them explicitly stated the reason for imposing such a constraint. One possible explanation is the numerical problem when using the logistic transformation for γ, as noted in Keane and Wasi (2013).

less (more) random answers to the choice questions in the treatment. Therefore, we can test our third hypothesis: providing static images, in addition to text descriptions, will reduce randomness in making choices, by examining the θ.

In addition, the treatment effect on the within-sample scale heterogeneity (i.e., the variance in scale) are revealed through λ, where a positive (negative) λ is associated with higher (lower) scale heterogeneity within each sample.

Given the above specification, the conditional probability of a respondent i choosing option q out of all Q options in a choice scenario n can be represented as

$$P_{iqn} = \frac{\exp[(\sigma_{in}\beta+\gamma\Gamma v_i+(1-\gamma)\sigma_{in}\Gamma v_i)A_{qn}]}{\sum_{q'=1}^{Q}\exp[(\sigma_{in}\beta+\gamma\Gamma v_i+(1-\gamma)\sigma_{in}\Gamma v_i)A_{q'n}]} \quad (4)$$

Hence, the simulated likelihood function, with R random draws on v_i and ε_{0n}, for I respondents choosing a sequence of N choice scenarios with Q options in each scenario is

$$\mathcal{L} = \sum_{i=1}^{I}\frac{1}{R}\sum_{r=1}^{R}\prod_{n=1}^{N}\prod_{q=1}^{Q}(P_{iqn})^{c} \quad (5)$$

where $c = 1$ if respondent i chooses q in choice scenario n, and $c = 0$ otherwise. Note that the pseudo panel nature of the data is taken into the estimation by the multiplication across all N scenarios.

The GMXL model is able to identify preferences with multimodal or heavy-tailed distributions of marginal utility, led by extreme behaviors such as near lexicographic preferences or highly random behavior at very low scale. It also has been empirically proved to outperforms simpler models such as the multinomial logit, the scaled multinomial logit and the mixed logit in terms of model fitness (Fiebig et al., 2010; Hess and Rose, 2012; Keane and Wasi, 2013). An issue that should be noted is the (in)separability between taste and scale parameters. With all parameters being random and allowed to be correlated, Hess and Rose (2012) show that the taste and scale parameters cannot be separately identified, because adding the random scale parameter is equivalent to using a model with more flexible distributions for taste parameters. In light of the assertion of Hess and Rose (2012), Czajkowski et al. (2016) use models with constraints on the correlation structure and demonstrate the use of introducing treatment-specific covariates into the mean and variance of scale parameter in GMXL.

In order to test our second hypothesis – providing static images in addition to text descriptions will better focus respondents' attention on the attributes – we apply the strategy proposed in Hess and Hensher (2010). The analysis derives posterior individual-specific coefficients following the "conditioning of individual tastes" strategy (Revelt and Train, 2000) to infer the level of attention respondents paid to attributes in choice questions. The individual-specific coefficients of marginal utility and its variance can be derived conditioned on the choices that a respondent makes for choice questions in choice experiments with multiple choice scenarios. The individual-specific coefficients conditioned on the sequence of observed choices, c_{iN}, of respondent i for all N choice scenarios are given by

$$\beta_i = \sum_{r=1}^{R} \frac{P(c_{iN}|A_{qN},\beta_r)}{\sum_{r=1}^{R} P(c_{iN}|A_{qN},\beta_r)} \beta_r \tag{6}$$

where β_r is a draw from the estimated means and covariance matrix of marginal utility, and A_{qN} is a matrix comprising the attributes of all N scenarios. The details of this can be found in chapter 11 of Train (2009). The individual-specific conditional variance of a particular element k, which measures the variability in β_i across all choice questions,[17] can be written accordingly as

$$Var(\beta_i^k) = \sum_{r=1}^{R} \frac{P\left(c_{iN}\middle|A_{qn},\beta_r\right)}{\sum_{r=1}^{R} P\left(c_{iN}\middle|A_{qn},\beta_r\right)} (\beta_r^k)^2 - \left[\sum_{r=1}^{R} \frac{P\left(c_{iN}\middle|A_{qn},\beta_r\right)}{\sum_{r=1}^{R} P\left(c_{iN}\middle|A_{qn},\beta_r\right)} \beta_r^k\right]^2 \tag{7}$$

Hess and Hensher (2010) suggest that a respondent can be labeled as ignoring a certain attribute if the conditional mean of their taste parameter (elements in β_i) has very high uncertainty, measured by the coefficients of variation for posterior individual-specific

[17] The intuition of the individual variance of β_i is that a respondent can have different marginal utility for each attribute in each choice question, because of effects such as learning or interactions with other attributes.

parameters. That is to say, given the individual-specific mean of β_i, the higher the variation of β_i across choice questions, the less attention a respondent pays to the attributes. They also suggest a criterion – a coefficient of variation greater than two – for allocating respondents to an attribute non-attending (ANA)[18] group.

Given our study design, for an individual i who takes survey treatment s, the utility function for option q of scenario n can be written as

$$U_{isqn} = \mathrm{B}_{1s}\mathrm{MedSpc}_{sqn} + \mathrm{B}_{2s}\mathrm{HiSpc}_{sqn} + \mathrm{B}_{3s}\mathrm{SmWtr}_{sqn} + \mathrm{B}_{4s}\mathrm{AwWtr}_{sqn} + \mathrm{B}_{5s}\mathrm{Mow0}_{sqn} + \mathrm{B}_{6s}\mathrm{Mow30}_{sqn} + \mathrm{B}_{7s}\mathrm{Mow70}_{sqn} + \mathrm{B}_{8s}\mathrm{LowPat}_{sqn} + \mathrm{B}_{9s}\mathrm{MedPat}_{sqn} + \beta_C\mathrm{Cost}_{qn} + \varepsilon_{iqn} \quad (8)$$

where B stands for $[\sigma_{in}\beta + \gamma\Gamma\nu_i + (1-\gamma)\sigma_{in}\Gamma\nu_i]$ in equation 1. All attributes other than the cost attribute are dummy coded, where a "low" level of diversity in plant species, "never" presence of water, "100%" area mowed, and "formal" level of pattern of planting are the basis for each attribute. See Table 1 for the names of the dummy coded variables. The above specification was first used to model responses from each survey version independently. To test differences between survey versions, joint estimations were conducted that combined two survey versions. In these joint estimations, the marginal utilities of the landscape attributes are allowed to vary between the versions, while the cost parameter is set as equal across samples to achieve identification.

In recent studies using a mixed logit or GMXL model, it is common practice to assume that all parameters are randomly distributed and correlated (Greene and Hensher, 2010; Hess and Hensher, 2010). In addition, some studies assume a log-normal distribution for the cost (or price) parameter so that the parameter will have its theoretically correct sign and

[18] Attribute non-attendance refers to the situation in which respondents ignore some of the attributes in choice scenarios and evaluate the utility of each scenario without considering the ignored attributes (Scarpa et al., 2009). ANA is one of the most studied issues in the recent literature of choice experiments, and many studies have shown that not accounting for ANA can lead to serious problems in the estimation results (Carlsson et al., 2010; Hess et al., 2013). Although this value is determined rather arbitrarily, we consider that it is reasonable. If given a C.V. that equals two, the probability of the mean being not significantly different from zero is approximately 30%.

avoid the problem of infinite moments for the distributions of willingness-to-pay (Hensher et al., 2012).[19] Given the relatively large number of parameters compared to the number of observations, we instead assume that the cost parameter is fixed, that all landscape attributes are normally distributed, and impose a restricted covariance matrix that only allows correlations between random parameters within each attribute and version (i.e., no correlation is allowed between levels of different attributes). Table A1 in the appendix shows an illustration of the covariance matrix.[20] It should be noted that, with the fixed cost parameter and restricted covariance matrix, a model without a random scale parameter would not fully capture the scale heterogeneity, which therefore grants the premise for separately identifying taste and scale heterogeneity.

In order to investigate the impacts of showing different information, samples from different versions are combined to estimate joint models. The first joint model combines the samples in the base and text-only versions, and examines the impacts of showing images in addition to text descriptions. The second joint model combines the samples in the base and image-only versions. When the version dummy variables are included, the estimates of θ and λ in equation 3 reveal the impacts of different information on scale. Empirically, the value of γ would not drastically change the estimation results. In the case where $\sigma_{in} = 1$, γ indeed has no impact on the utility function. Therefore, we estimate Type I GMXL models with $\gamma = 1$ to reduce unspecified correlations between the willingness-to-pay estimates. Note that given the current setting, σ_{in} reduces to σ_i, since no scale heterogeneity resulting from different choice scenarios is specified.

[19] An issue with using a log-normally distributed cost parameter is that, with its fat-tail feature, a relatively large portion of people will have very low marginal utility for cost resulting in very large willingness-to-pay (WTP) values. Although models in WTP space can largely alleviate this problem, estimating a GMXL model in WTP space is empirically impossible because the scale parameter cannot be separated from the cost parameter (Hole and Kolstad, 2012).

[20] We estimated a GMXL with unrestricted covariance matrix (i.e., allowing every parameters to be correlated with each other within treatment), but the model did not converge. We examined the results with mixed logit model in WTP space with all parameters being normally distributed and fully correlated. The WTP estimates were qualitatively similar to those derived from our GMXL model.

All GMXL model estimations are executed using a Matlab package, Models for Discrete Choice Experiments (Czajkowski, 2018), using 1,000 Halton draws. We used several sets of starting values to ensure the identification and convergence (Czajkowski et al., 2015).

Empirical Results

For the sake of brevity, in this section, diversity in plant species is referred to as "diversity;" the presence of water is referred to as "water;" the percentage of green spaces mowed is referred to as "area mowed;" and the pattern of plantings is referred to as "planting patterns." Table 3 presents the results of the GMXL models using each of the three samples.[21] In table 4, we calculate the corresponding conditional willingness-to-pay (WTP) for the landscape attributes using the individual-specific coefficients (see equation 6).

High diversity is viewed as an amenity, although its estimate in the image-only version is of marginal significance. The WTPs for high diversity range from $20 to $66 in terms of the annual homeowner association fee. Respondents also showed a positive attitude towards medium diversity when described textually.[22] The significant estimates of permanent water across all three models indicate that people are worried about whether the designed green spaces always have a pool of water, and the WTPs for avoiding such a design are between $66 and $118. Even intermittent water, which drains within one day after a rainstorm, is unfavorable when illustrated using the images. In the base treatment, people prefer medium levels of area mowed (30% and 70%) to the two extreme levels, while the preferences for area mowed show fairly different patterns across the three samples. Lastly, the respondents generally view the planting pattern indifferently.

[21] The estimated covariance matrixes of all models are available upon request.

[22] The finding that people are in favor of diversity is consistent with the majority of studies that look at the preferences for biodiversity (e.g., Birol et al., 2006; Christie et al., 2006; Hanley et al., 1998).

Overall, respondents consistently placed positive values on diversity and negative values on water, while the level of area mowed and whether the planting pattern was formal were of less concern. We also identify considerable taste heterogeneity through the significant standard deviation coefficients for all taste parameters except those of the planting pattern. Therefore, the results also indicate that a portion of respondents prefer lower diversity or more water.

Impacts of Information Treatments on Taste Parameters

In the model results presented in Table 5, the samples from the base and text-only versions are combined to estimate the GMXL model with the inclusion of covariates on scale parameters. The results shown in Table 6 use samples from the base and image-only versions.[23]

We find evidence to support our hypothesis claiming people's taste will change with the provision of static visualizations. When both images and text descriptions are used to describe choice scenarios, respondents exhibit significantly stronger preferences for diversity and water attributes. For example, the marginal utility for higher diversity is at least three times greater in the base version than in the image only or text only version. A similar pattern is identified in the preferences for water: the coefficients in the base version samples are much greater in magnitude than those in the other two samples. These patterns are also confirmed in the WTPs reported in Table 4. This finding is in contrast to that in Bateman et al. (2009), who found their estimates were larger in magnitude when the scenarios are presented using only textual descriptions.

In Table 6, the significant estimates for diversity and water in the image-only version samples prove that, even absent textual descriptions, a landscape with high diversity is visually preferable, and water is also visually salient but viewed as a disamenity. Observing

[23] Tests of the equality of the mean values of the coefficients are presented in Table A2 in the appendix.

the similar WTPs (Table 4) and the coefficients of diversity and water in the text-only and image-only samples (Tables 5 and 6) suggests that images may be as effective as text descriptions at conveying these two landscape attributes with high visual salience. A noticeable impact of showing images on the preferences for area mowed is that the visual stimuli ease people's concerns for 0% area mowed in text-only samples [24], when both images and text depictions are offered. In the image-only treatment, on average none of the level of area mowed is visually preferred to one another, while people still have very heterogeneous preferences for each level. Lastly, the coefficients on planting pattern do not significantly differ across samples.

The relative standard deviations (standard deviation divided by mean) for the diversity and water attributes are generally greater in text-only and image-only versions than in the base version. Our conjecture is that respondents could better understand what the attributes and their levels exactly mean when shown a more complete information set, and they could thus make their decisions based on the information that the researchers intended to provide, rather than infer their own ideas.

In summary, people's preferences regarding landscape attributes can be largely underestimated without proper information that describes the attributes, and both visual representations and textual descriptions can help respondents better understand what they are being asked to value. This indicates greater overall "evaluability" of the information with visual representation (Bateman et al., 2009; Hsee, 1996). Still, we note some caveats regarding the use of visual representations. For example, respondents could have read more information from images than that which the researchers intended to offer. First, respondents might see differences between pictures that are not intended, such as

[24] This may be attributed to people's concerns such as the green spaces being not accessible or becoming overgrown. Some locations legislate against leaving areas to grow, as this is thought to symbolize a lack of care and municipalities believe that such areas harbor vermin – rats, snakes, and undesirable insects.

differences in terms of whether the green spaces are well maintained and if the water area is accessible.[25] Second, unintended differences can be created by the interactions between landscape attributes (Dramstad et al., 2006), which is related to the discussion in footnote 10 regarding how the change in housing density could result in some randomness in the images. Lastly, as previously pointed out, static images might not be able to adequately capture the dynamic aspects of attributes, such as how long the water stays.

Impacts of Information Treatments on Attribute Non-attendance (ANA)

Individual-specific means and standard deviations were derived based on the three GMXL models presented in Table 3. Table 7 shows the percentages of respondents in the ANA group, whose posterior individual-specific coefficient of variation (C.V.) greater than two.

The rates of ANA for any specific attribute were generally the lowest for respondents to the base version survey. In particular, ANA rates for high diversity and permanent water in the base version sample were less than half those for the text-only version samples. This suggests that people are twice as likely to ignore the diversity and water attributes when the choice questions are not described with images.[26] In addition, the means and medians of the individual-specific coefficients of variation, also reported in Table 7, are consistently higher in the base version sample. We therefore find evidence to support our second hypothesis – providing images better focuses respondents' attention on the attributes. Such impacts from the provision of images are attributable to the notion that visualizations can increase people's curiosity and enhance the clarity of their stated preferences (Barroso et al., 2012).

[25] These two attributes are associated with the aspect of engagement in the landscape. Humans by nature like environments that offer opportunities for meaningful engagement, so a landscape with no apparent sign of human activity was likely to be unfavorable (Kaplan and Kaplan, 1989; Zacharias, 2001).

[26] We focus on the rates of ANA for high diversity and permanent presence of water because the marginal utility estimates for these attributes are mostly statistically significant across all treatments. The approach we employ tends to identify more respondents who ignore an attribute when the conditional mean of a posterior distribution is closer to zero. In addition, ANA rates can also largely change when different base levels for each attribute are used in dummy coding.

Impacts of Information Treatments on Scale Parameters

In Tables 5 and 6, the two positively significant estimates of θ indicate that the mean scale is lower when the respondents are provided with both images and text descriptions of choice scenarios as compared to image-only or text-only. On average, the base sample responses are more random than the responses in the text-only and image-only samples.[27] We therefore find the opposite of our third hypothesis on the randomness of making choices.

Although this finding is in contrast to those found in previous stated preference studies using dynamic visualizations (Bateman et al., 2009; Matthews et al. 2017), it is instead in line with Townsend and Kahn (2014) – using visual representation can make people use less systematic cognitive processes when making decisions. Essentially, such an impact can be created by increasing the complexity of choice tasks. Extra images or text descriptions make the information set both more "complete" and more "diverse." We argue that complete information can facilitate the decision-making process, while diverse information requires more cognitive processing. Therefore, the lower mean scale parameter when extra information is supplied can be considered because, on average, the effect of more diverse information outweighs the effect of higher information completeness. Combining the findings of more attention to attributes and more random choices with visualizations suggests a potential information overload problem.

One explanation for why complexity makes choices more random is that people might not be able to precisely evaluate the total utility of each option for each choice when the complexity of their choice task increases. This is the situation where people find making a choice daunting because they know too much and they finally just go with their gut. Another explanation is that humans tend to use automatic processes in cognitive processing to process visual information, while automatic processes are less systematic than their

[27] We again acknowledge the assertion of Hess and Rose (2012) regarding inseparability of scale and coefficient heterogeneity. However, regardless whether the difference is identified as a lower scale parameter or as lower marginal utilities, both interpretations indicate more "noisy" choices.

counterpart, controlled processes, which are used to handle textual information (Townsend and Kahn, 2014).

Another rather different interpretation of this finding from a more economic perspective is that people behave more extremely when information is sparse or insufficient. This interpretation is in agreement with Fiebig et al. (2010), who suggest that people's behavior is less extreme when the choice tasks are more complex. In the context of the current study, when only text descriptions are provided, some "nature lovers" would make their choices solely based on the attribute of diversity (e.g., always go with the option with higher diversity) and ignore other attributes. That is, they have lexicographical preference. When images are provided, they would realize that they also wanted to see water near the houses, which meant that they now needed to make a tradeoff. van der Wal et al. (2014) also found that the provision of new information of which people were not formerly aware could make people less likely to uphold extreme behavior in their landscape preferences. A similar story can be found between the image-only and base version samples. People might not realize that the images actually illustrate four landscape attributes, so their decisions were made based on the fewer characteristics that they have observed. Once the text descriptions are attached, they need more cognitive effort to consider more attributes and the exact levels, so the decision process is not as simple as when only images are provided.

Among the two studies that have investigated the impact of (non-visual) information provision on the scale parameters, Czajkowski et al. (2016) find that giving a more complete information set regarding biodiversity conservation programs made respondents give more predictable choices (i.e., higher mean scale parameter). However, Lusk et al. (2008) find the opposite in their experiment when providing more information about the benefits of grass-fed beef. The treatment performed in the latter study was directly linked to a certain attribute (grass-fed or not), so the complexity of the choice tasks could be increased when the extra information was new to the respondents or when an attribute was previously ignored by some respondents in the case where the extra information was absent. This explanation is consistent with that proposed in the previous paragraph. However, the treatment carried out in Czajkowski et al. (2016) involved much more than just adding information about the attributes. The alternate version of their survey was rewritten according to the feedback of a group of stakeholders and intended to provide a more comprehensive context for respondents. The revised survey emphasized the positive impacts and consequences of the program in which they were interested, so respondents' preferences could be somewhat guided in favor of biodiversity conservation. It should be no surprise that their conclusion regarding the impact of the scale is considerably different from those of the current paper and of Lusk et al. (2008).

Turning to the variances of scale parameters, the significant τ in Table 5 indicates that the scale parameter is heterogeneous across respondents. Although the τ in Table 6 is not significant, its magnitude is comparable to that in Table 5. The insignificant estimates of λ in Tables 5 and 6 suggest that the variability in scale parameter is not affected when the respondents are provided with different information.

It is important to clarify the difference between preference uncertainty for a certain attribute, which can be measured by the coefficient of variation of marginal utility, and randomness when making choices,[28] which is measured by a scale parameter. The preference uncertainty for a certain attribute shows how well people know their value for an attribute. This study proposes that people will be less likely to ignore attributes and better informed on the condition of the attributes when more information is provided, thus the preference uncertainty for a certain attribute can be lower. However, the randomness of making choices across all choice scenarios can still be higher due to the greater information received, as discussed in the previous subsection.

[28] This dimension is also interpreted as the certainty of their choices, which refers to whether people believe that they have picked the preferred option (Kiani et al., 2014).

Conclusions

Employing visual representations to demonstrate choice scenarios in stated preference surveys is a long-standing practice, particularly when the research questions are associated with landscapes. However, the impacts of such practices, using static images in particular, on preference were largely unknown, despite the fact that the policy implications can be greatly influenced by the use of visualizations. This paper conducts a choice experiment looking at citizen preferences on landscape attributes of green infrastructure. With three survey treatments (images plus text, text-only, and image-only), we provide a comprehensive examination of the impacts of using static visualizations on marginal utility estimates, randomness of responses, and attention to questions. The findings are also valuable for better GI designs.

Our findings indicate that providing static images changes people's taste and better focuses respondents' attention on the attributes; however, it increases randomness in making choices. When respondents were presented with both text and images, they exhibited stronger preferences for visually salient landscape attributes and were less likely to ignore those attributes than when presented with either only images or only text. Although people pay more attention to individual attributes in choice questions with the provision of visual representations, the extra visual stimuli also increase choice randomness, as measured by the scale parameter. This may be attributed to the cognitive overload caused by paying more attention to the questions and the need to process more information.

In terms of the preferences for GI landscape attributes, the results show that most people prefer richer variety in plant species but dislike having water present in green spaces. Given the assumption that people can better understand planned landscape changes when using images and text together, the estimates for the taste parameters of the two attributes just mentioned, which have high visual salience, are significantly underestimated (in magnitude) when not illustrated with both images and text. In addition to GI's hydrological benefits, homeowners' acceptance of GI also depends on whether they can understand and accept the potential landscape changes and associated environmental benefits that result from introducing GI to their neighborhoods. Therefore, it is necessary to provide both images and text descriptions in choice experiments to accurately capture the values that people place on visually salient attributes.

This paper concludes that including visual representations is preferable when using choice experiments to value landscape-related issues. In addition, accurate visualizations are essential to prevent biased results, given the findings that responses are susceptible to the method used to convey questions. Furthermore, researchers should be aware that the additional information contained in visual images may increase decision complexity, resulting in less predictable responses. Future research should investigate the optimal amount of information when describing landscape attributes in choice experiments (i.e., the amount of information that can help respondents best understand the questions without overwhelming them). In particular, studies have shown that people's experience and familiarity with a landscape are related to their corresponding preferences (Rogge et al., 2007; Scott et al., 2009). Incorporating experience and familiarity with a landscape and associated amenities into the analysis is a promising line for future research to complement the investigation of how different visual information can affect people's choices.

References

Barroso, Filipe Lucas, Teresa Pinto-Correia, Isabel L. Ramos, Diana Surová, and Helena Menezes. 2012. Dealing with landscape fuzziness in user preference studies: Photo-based questionnaires in the Mediterranean context. *Landscape and Urban Planning* 104 (3–4): 329–342

Bateman, Ian J., Brett H. Day, Andrew P. Jones, and Simon Jude. 2009. Reducing gain–loss asymmetry: a virtual reality choice experiment valuing land use change. *Journal of Environmental Economics and Management* 58(1): 106–118

Beharry-Borg, Nesha, and Riccardo Scarpa. 2010. Valuing quality changes in Caribbean coastal waters for heterogeneous beach visitors. *Ecological Economics* 69(5): 1124–1139

Berninger, Kati, Wiktor Adamowicz, Daniel Kneeshaw, and Christian Messier. 2010. Sustainable forest management preferences of interest groups in three regions with different levels of industrial forestry: An exploratory attribute-based choice experiment. *Environmental management* 46(1): 117–133

Birol, Ekin, Katia Karousakis, and Phoebe Koundouri. 2006. Using a choice experiment to account for preference heterogeneity in wetland attributes: the case of Cheimaditida wetland in Greece. *Ecological Economics* 60(1): 145–156

Campbell, Danny, Marco Boeri, Edel Doherty, and W. George Hutchinson. 2015. Learning, fatigue and preference formation in discrete choice experiments. *Journal of Economic Behavior & Organization* 119: 345-363.

Carlsson, Fredrik, Mitesh Kataria, and Elina Lampi. 2010. Dealing with ignored attributes in choice experiments on valuation of Sweden's environmental quality objectives. *Environmental and Resource Economics* 47(1): 65–89

Chernev, Alexander, Ulf Böckenholt, Joseph Goodman. 2015. Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology* 25(2): 333–358

Christie, Mike, Nick Hanley, John Warren, Kevin Murphy, Robert Wright, and Tony Hyde 2006. Valuing the diversity of biodiversity. *Ecological Economics* 58(2): 304–317

Czajkowski, Mikołaj. 2018. Models for discrete choice experiments. Https://github.com/czaj/DCE

Czajkowski, Mikołaj, Nick Hanley, and Jacob LaRiviere. 2015. The effects of experience on preferences: theory and empirics for environmental public goods. *American Journal of*

Agricultural Economics 97(1): 333–351

Czajkowski, Mikołaj, Nick Hanley, and Jacob LaRiviere. 2016. Controlling for the effects of information in a public goods discrete choice model. *Environmental and Resource Economics* 63(3): 523–544

De Valck, Jeremy, Pieter Vlaeminck, Steven Broekx, Inge Liekens, Joris Aertsens, Wendy Chen, and Liesbet Vranken. 2014. Benefits of clearing forest plantations to restore nature? Evidence from a discrete choice experiment in Flanders, Belgium. *Landscape and Urban Planning* 125: 65–75

del Carmen Torres-Sibille, Ana, Vicente-Agustín Cloquell-Ballester, Víctor-Andrés Cloquell-Ballester, and Miguel Ángel Artacho Ramírez. 2009. Aesthetic impact assessment of solar power plants: An objective and a subjective approach. *Renewable and Sustainable Energy Reviews* 13(5): 986–999

Dimitropoulos, Alexandros, and Andreas Kontoleon. 2009. Assessing the determinants of local acceptability of wind-farm investment: A choice experiment in the Greek Aegean Islands. Energy policy 37(5): 1842–1854

Dramstad, W. E., M. Sundli Tveit, W.J. Fjellstad, and G.L.A. Fry. 2006. Relationships between visual landscape preferences and map-based indicators of landscape structure. *Landscape and Urban Planning* 78(4): 465–474

Elsasser, Peter, Hermann Englert, and Jacqueline Hamilton. 2010. Landscape benefits of a forest conversion pro-gramme in North East Germany: results of a choice experiment. *Annals of Forest Research* 53(1): 37–50

Fiebig, Denzil G., Michael P. Keane, Jordan Louviere, and Nada Wasi. 2010. The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. *Marketing Science* 29(3): 393–421

Greene, William H., and David A. Hensher. 2010. Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models. *Transportation* 37(3): 413–428

Haider, Wolfgang, Anderson, D.A., Daniel, T.C., Louviere, J.J., Orland, B., and Williams, M. 1998. Combining calibrated digital imagery and discrete choice experiments: An application to remote tourism in Northern Ontario. Shaping Tomorrow's North – Proceedings of an International Conference, Thunder Bay, ON, 257– 278.

Haider, Wolfgang, and Len Hunt. 2002. Visual aesthetic quality of northern Ontario's forested shorelines. *Environmental Management*, 29(3), 324-334.

Hanley, Nick, Robert E. Wright, and Vic Adamowicz. 1998. Using choice experiments to value the environment. *Environmental and Resource Economics* 11(3–4): 413–428

Harline, Carl E., and Mark W. Burris. 2014. Impact of traffic images on route choice and the value of time estimates in stated preference surveys. *Transportation Research Record: Journal of the Transportation Research Board* (2405): 24–32

Hensher, David A., Matthew J. Beck, and John M. Rose. 2011. Accounting for preference and scale heterogeneity in establishing whether it matters who is interviewed to reveal household automobile purchase preferences. *Environmental and Resource Economics* 49(1): 1–22

Hensher, David A., Jordan Louviere, and Joffre Swait. 1998. Combining sources of preference data. *Journal of Econometrics* 89(1): 197–221

Hensher, David A., John M. Rose, and Zheng Li. 2012. Does the choice model method and/or the data matter? *Transportation* 39(2): 351–385

Hess, Stephane, and David A. Hensher. 2010. Using conditioning on observed choices to retrieve individual-specific attribute processing strategies. *Transportation Research Part B: Methodological* 44(6): 781–790

Hess, Stephane, and John M. Rose. 2012. Can scale and coefficient heterogeneity be separated in random coefficients models? *Transportation* 39(6): 1225–1239

Hess, Stephane, Amanda Stathopoulos, Danny Campbell, Vikki O'Neill, and Sebastian Caussade 2013. It's not that I don't care, I just don't care very much: confounding between attribute non-attendance and taste heterogeneity. *Transportation* 40(3): 583–607

Hill, Dawn, Terry C. Daniel. 2007. Foundations for an ecological aesthetic: Can information alter landscape preferences? *Society & Natural Resources* 21(1): 34–49

Hole, Arne Risa, and Julie Riise Kolstad. 2012. Mixed logit estimation of willingness to pay distributions: a comparison of models in preference and WTP space using data from a health-related choice experiment. *Empirical Economics* 42(2), 445-469.

Holmes, Thomas P., and Wiktor L. Adamowicz. 2003. Attribute-based methods. In *A Primer on Nonmarket Valuation*. Springer Netherlands.

Hsee, Christopher K. 1996. The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes* 67(3): 247-257

Hull IV, R. Bruce, and Grant R.B. Revell. 1989. Issues in sampling landscapes for visual quality assessments. *Landscape and Urban Planning* 17(4): 323–330

Iglesias, Paula, Margarita Greene, and Juan de Dios Ortúzar. 2013. On the perception of safety in low income neighbourhoods: using digital images in a stated choice experiment. *Choice Modelling: The State of the Art and the State of Practice*, 193-210.

Kaplan, Rachel, and Stephan Kaplan. 1989. *The Experience of Nature: A Psychological Perspective*. Cambridge University Press.

Keane, Michael, and Nada Wasi. 2013. Comparing alternative models of heterogeneity in consumer choice behavior. *Journal of Applied Econometrics* 28(6): 1018–1045

Kiani, Roozbeh, Leah Corthell, and Michael N. Shadlen. 2014. Choice certainty is informed by both evidence and decision time. *Neuron* 84(6): 1329–1342

Londoño Cadavid, Catalina, and Amy W. Ando. 2013. Valuing preferences over stormwater management outcomes including improved hydrologic function. *Water Resources Research* 49(7): 4114–4125

Louviere, Jordan J., Robert J. Meyer, David S. Bunch, Richard Carson, Benedict Dellaert, W. Michael Hanemann, David Hensher, and Julie Irwin. 1999. Combining sources of preference data for modeling complex decision processes. *Marketing Letters* 10(3): 205–217

Lusk, Jayson L., Deacue Fields, and Walt Prevatt. 2008. An incentive compatible conjoint ranking mechanism. *American Journal of Agricultural Economics* 90(2): 487–498

Manning, Robert E., and Wayne A. Freimund. 2004. Use of visual research methods to measure standards of quality for parks and outdoor recreation. *Journal of Leisure Research* 36(4): 557–579

Mathews, Kristy E., Miranda L. Freeman, and William H. Desvousges. 2006. How and how much? The role of information in stated choice questionnaires, In *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice*. Springer.

Matthews, Yvonne, Ricardo Scarpa, and Dan Marsh. 2017. Using virtual environments to improve the realism of choice experiments: A case study about coastal erosion management. *Journal of Environmental Economics and Management*, 81: 193–208

Natori, Yoji, and Richard Chenoweth. 2008. Differences in rural landscape perceptions and preferences between farmers and naturalists. *Journal of Environmental Psychology* 28(3):

250–267
Patterson, Zachary, Javad Mostofi Darbani, Ali Rezaei, John Zacharias, and Ali Yazdizadeh. 2017. Comparing text-only and virtual reality discrete choice experiments of neighbourhood choice. *Landscape and Urban Planning* 157: 63–74
Revelt, David, and Kenneth Train. 2000. Customer-specific taste parameters and Mixed Logit: Households' choice of electricity supplier. *Department of Economics, UCB*. Unpublished Manuscript
Rid, Wolfgang, Wolfgang Haider, Andrea Ryffel, and Ben Beardmore. 2018. Visualisations in Choice Experiments: Comparing 3D Film-sequences and Still-images to Analyse Housing Development Alternatives. *Ecological Economics*, *146*, 203-217.
Rizzi, Luis I., Juan Pablo Limonado, and Seiji S.C. Steimetz. 2012. The impact of traffic images on travel time valuation in stated-preference choice experiments. *Transportmetrica* 8(6): 427–442
Rogge, Elke, Frank Nevens, and Hubert Gulinck. 2007. Perception of rural landscapes in Flanders: Looking beyond aesthetics. *Landscape and Urban Planning* 82(4): 159–174
Ryffel, Andrea, Wolfgang Rid, Adrienne Grêt-Regamey. 2014. Land use trade-offs for flood protection: A choice experiment with visualizations. *Ecosystem Services* 10: 111–123.
Scarpa, Riccardo, Danny Campbell, and W. George Hutchinson. 2007. Benefit estimates for landscape improvements: sequential Bayesian design and respondents' rationality in a choice experiment. *Land Economics* 83(4): 617–634
Scarpa, Riccardo, Timothy J. Gilbride, Danny Campbell, and David A. Hensher. 2009. Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European review of agricultural economics* 36(2): 151–174
Scott, Alister, Claudia Carter, Katrina Brown, and Vicki White. 2009. 'Seeing is not everything': Exploring the landscape experiences of different publics. *Landscape Research* 34(4): 397–424
Schäffler, Alexis, and Mark Swilling. (2013). Valuing green infrastructure in an urban environment under pressure—The Johannesburg case. *Ecological Economics* 86: 246–257
Townsend, Claudia, and Barbara E. Kahn. 2014. The "visual preference heuristic": the influence of visual versus verbal depiction on assortment processing, perceived variety, and choice overload. *Journal of Consumer Research* 40(5): 993–1015
Train, Kenneth. 2009. Drawing from densities. In *Discrete Choice Methods with Simulation, 2nd*

Ed. Cambridge University Press.

Tzoulas, Konstantinos, Kalevi Korpela, Stephen Venn, Vesa Yli-Pelkonen, Aleksandra Kaźmierczak, Jari Niemela, and Philip James. 2007. Promoting ecosystem and human health in urban areas using Green Infrastructure: A literature review. *Landscape and Urban Planning* 81(3): 167–178

United States Environmental Protection Agency (USEPA). 2016. What is Green Infrastructure? https://www.epa.gov/green-infrastructure/what-green-infrastructure

van der Wal, René, David Miller, Justin Irvine, Stefano Fiorini, Arjun Amar, Steven Yearley, Robin Gill, and Norman Dandy. 2014. The influence of information provision on people's landscape preferences: A case study on understorey vegetation of deer-browsed woodlands. *Landscape and Urban Planning* 124: 129–139

Williams, Kathryn J.H., Rebecca M. Ford, Ian D. Bishop, Daniel Loiterton, John Hickey. 2007. Realism and selectivity in data-driven visualisations: A process for developing viewer-oriented landscape surrogates. *Landscape and Urban Planning* 81(3), 213–224

Zacharias, John. 2001. Path choice and visual stimuli: signs of human activity and architecture. *Journal of Environmental Psychology* 21(4): 341–352

Tables

Table 1. Landscape Attributes and Levels of Green Infrastructure

Attributes	Variety in Plant Species (Diversity)	Presence of Water (Water)	Percentage of Mowed Area (Area Mowed)	Pattern of Plantings (Planting Pattern)	Cost
Levels	High (HiSpc) Medium (MedSpc) Low*	Always (AwWtr) Sometimes (SmWtr) Never*	0% (Mow0) 30% (Mow30) 70% (Mow70) 100%*	Informal (LowPat) Intermediate (MedPat) Formal*	\$110 \$100 \$90 \$55 \$50 \$45 \$0

*: Base levels.

Note: Variable names used in equation 10 are in parenthesis.

Table 2. Socio-Demographic Profiles of Respondents by Survey Version

	Base (n = 159)	Text-only (n = 170)	Image-only (n = 170)
Mean Age	45.91 [50.51]	45.74 [47.81]	46.21 [50.54]
Female	55.91% [63.52%]	53.96% [68.24%]	50.01% [65.88%]
College Educated	42.91% [57.86%]	39.93% [53.53%]	38.82% [60.00%]
Live in Single Family Houses	56.10% [69.81%]	58.72% [62.94%]	55.77% [66.47%]
Percentage choosing low housing density neighborhood	81.13% [83.69%]	82.94% [80.23%]	77.06% [74.81%]

Note: unweighted summary statistics are in brackets.

Table 3. Results of the GMXL Models for Each Single Version

		Base		Text-Only		Image-Only	
Attribute	Level	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
		Mean					
Diversity	Medium	1.3999***	0.2751	0.5417***	0.1846	0.1027	0.3390
	High	1.7415***	0.3570	0.8127***	0.2158	0.6462*	0.3514
Water	Sometimes	-0.7290***	0.1624	-0.1254	0.1681	-0.4553***	0.1104
	Always	-3.1396***	0.3617	-2.4801***	0.3429	-2.1697***	0.3699
Mowed	0%	0.3627	0.3499	-1.4790***	0.3578	-0.2273	0.3699
	30%	0.6117**	0.2527	0.2693	0.2643	0.0358	0.2612
	70%	0.9643***	0.2865	0.6080**	0.2492	-0.2981	0.3174
Planting	Medium	-0.1403	0.1631	0.1801	0.1763	-0.3298*	0.1741
Pattern	Low	0.1915	0.2509	-0.4227*	0.2499	-0.2986	0.1835
Cost		-0.0266***	0.0050	-0.0376***	0.0064	-0.0319***	0.0055
		Standard Deviations					
Diversity	Medium	1.6432***	0.2361	1.3119***	0.2492	2.4934***	0.2795
	High	2.4346***	0.3269	1.5469***	0.2203	2.2105***	0.2844
Water	Sometimes	1.1912***	0.2150	1.4400***	0.2151	0.5743***	0.1831
	Always	2.8573***	0.3523	2.7042***	0.2801	2.8129***	0.3231
Mowed	0%	1.7704***	0.3292	2.6765***	0.3468	0.6380	0.4762
	30%	1.1559***	0.3032	1.7347***	0.2876	0.7591*	0.4093
	70%	0.7815**	0.3502	1.0455***	0.3470	1.0382**	0.4379
Planting	Medium	0.3281	0.2427	0.4398	0.2639	0.5538***	0.1880
Pattern	Low	0.6642*	0.3400	1.2799***	0.2553	1.1019*	0.6235
τ		0.6870	0.5966	1.0387***	0.3552	1.1497***	0.5169
Log-Likelihood Ratio		-845.87		-1006.8		-999.19	
Observations		1908		2040		2040	
Parameters		26		26		26	

*, **, ***: Significant at the 10%, 5%, and 1% levels, respectively.

Table 4. Willingness-to-pay for the Landscape Attributes of Green Infrastructure

		Base		Text-only		Image-only	
Attribute	Level	Mean	S.D.	Mean	S.D.	Mean	S.D.
Diversity	Medium	52.65*	61.80*	14.42*	34.93*	3.22	78.10*
	High	65.50*	91.57*	21.63*	41.18*	20.24	69.23*
Water	Sometimes	-27.42*	44.80*	-3.34	38.34*	-14.26*	17.99*
	Always	-118.09*	107.46*	-66.02*	71.99*	-67.96*	88.10*
Mowed	0%	13.64	66.59*	-39.37*	71.25*	-7.12	19.98
	30%	23.01*	43.48*	7.17	46.18*	1.12	23.77
	70%	36.27*	29.39*	16.19*	27.83*	-9.34	32.52*
Planting	Medium	-5.28	12.34	4.80	11.71	-10.33	17.34*
Pattern	Low	7.20	24.98	-11.25	34.07*	-9.35	34.51

*: Significant at the 5% level.

Table 5. Results of the GMXL Model with the Base and Text-only Version Samples

		Base		Text-only	
Attribute	Level	Coeff.	S.E.	Coeff.	S.E.
		Mean		Mean	
Diversity	Medium	1.3069***	0.2533	0.3321**	0.1633
	High	1.4695***	0.3231	0.2872*	0.1489
Water	Sometimes	-0.6728***	0.1793	-0.1004	0.1001
	Always	-3.0862***	0.3319	-1.6584***	0.1749
Mowed	0%	0.2443	0.3634	-1.0908***	0.2854
	30%	0.5830**	0.2353	0.0881	0.1963
	70%	1.0127***	0.3452	0.2725	0.1770
Planting	Medium	-0.0819	0.1703	0.1397	0.1144
Pattern	Low	0.1173	0.3156	-0.2538**	0.1183
Cost		-0.0250***	0.0034	Equal across samples	
		S.D.		S.D.	
Diversity	Medium	1.7222***	0.2465	0.9115***	0.1608
	High	2.5421***	0.3171	1.0135***	0.1247
Water	Sometimes	1.1287***	0.2210	0.9736***	0.1427
	Always	2.4316***	0.3224	1.8128***	0.1877
Mowed	0%	1.6768***	0.3576	1.7778***	0.1988
	30%	1.3008***	0.3204	1.2342***	0.1569
	70%	0.9953**	0.4184	0.7581***	0.1755
Planting	Medium	0.5122	0.3848	0.2732*	0.1444
Pattern	Low	0.7296*	0.3989	0.7499***	0.1786
θ		0.3419***	0.1202		
τ		0.9990**	0.4729		
λ		-0.2740	0.6239		
Log-Likelihood Ratio		-1858.9			
Observations		3948			
Parameters		52			

*, **, ***: Significant at the 10%, 5%, and 1% levels, respectively.

Table 6. Results of the GMXL Model with the Base and Image-only Version Samples

		Base		Image-only	
Attribute	Level	Coeff.	S.E.	Coeff.	S.E.
		Mean		Mean	
Diversity	Medium	1.5228***	0.2961	0.1717	0.1673
	High	1.9048***	0.4127	0.6060***	0.1640
Water	Sometimes	-0.7569***	0.1847	-0.3902***	0.0846
	Always	-3.3878***	0.3352	-1.8633***	0.2298
Mowed	0%	0.3553	0.4075	-0.2606	0.2135
	30%	0.6395**	0.2601	-0.0054	0.1759
	70%	1.0190***	0.3220	-0.2811	0.1893
Planting	Medium	-0.1279	0.1946	-0.2579**	0.1245
Pattern	Low	0.1747	0.3385	-0.2682**	0.1357
Cost		-0.0276***	0.0029	Equal across samples	
		S.D.		S.D.	
Diversity	Medium	1.7226***	0.2736	1.9654***	0.1762
	High	2.7689***	0.3821	1.8411***	0.1910
Water	Sometimes	1.2752***	0.2200	0.6471***	0.1432
	Always	2.5475***	0.3275	2.5903***	0.2247
Mowed	0%	1.9796***	0.3157	0.2890*	0.1734
	30%	1.3418***	0.3372	0.5256***	0.2033
	70%	0.9208**	0.3876	0.7173***	0.2064
Planting	Medium	0.3867	0.2660	0.6158***	0.1302
Pattern	Low	0.7857*	0.4053	0.5222***	0.1892
θ		0.4751***	0.1428		
τ		0.8638	0.6606		
λ		0.2932	0.6700		
Log-Likelihood Ratio		-1812.5			
Observations		3948			
Parameters		52			

*, **, ***: Significant at the 10%, 5%, and 1% levels, respectively.

Table 7. Rates of Attribute Non-Attendance on Diversity of Plant Species and Presence of Water

	Percentage of ANA			Average Coefficient of Variation			Median Coefficient of Variation		
	Base	Text -only	Image -only	Base	Text -only	Image -only	Base	Text -only	Image-only
Diversity									
Medium	8.81	14.47	4.12	3.932	5.578	7.026	0.711	1.341	0.786
High	7.75	15.75	10.59	1.347	2.799	1.825	0.778	1.555	0.909
Water									
Sometimes	4.40	13.21	14.12	1.257	2.718	2.660	0.587	1.146	0.836
Always	3.14	6.29	9.41	1.510	1.698	3.102	0.441	0.675	0.589

Figures

Figure 1A. Example Choice Question in the Base Version (with Low Housing Density)

			
Neighborhood A		Neighborhood B	
Diversity in Species	Low	Diversity in Species	High
Standing Water	Never	Standing Water	Always
Percentage of Green Space Mowed	30%	Percentage of Green Space Mowed	70%
Pattern of Plantings	Formal	Pattern of Plantings	Intermediate
Annual Homeowner Association Fee	$100	Annual Homeowner Association Fee	$50
○		○	

Figure 1B. Example Choice Question in the Base Version (with Medium Housing Density)

Neighborhood A		Neighborhood B	
Diversity in Species	Low	Diversity in Species	High
Standing Water	Never	Standing Water	Always
Percentage of Green Space Mowed	30%	Percentage of Green Space Mowed	70%
Pattern of Plantings	Formal	Pattern of Plantings	Intermediate
Annual Homeowner Association Fee	$100	Annual Homeowner Association Fee	$50
○		○	

Figure 1C. Example Choice Question in the Base Version (with High Housing Density)

			
Neighborhood A		Neighborhood B	
Diversity in Species	Low	Diversity in Species	High
Standing Water	Never	Standing Water	Always
Percentage of Green Space Mowed	30%	Percentage of Green Space Mowed	70%
Pattern of Plantings	Formal	Pattern of Plantings	Intermediate
Annual Homeowner Association Fee	$100	Annual Homeowner Association Fee	$50
○		○	

Appendix

Table A1. Covariance Matrix Structure

	Att1.1[a]	Att1.2[a]	Att2.1[a]	Att2.2[a]	…	Att1.1[b]	Att1.2[b]	Att2.1[b]	Att2.2[b]	…	Cost
Att1.1[a]	#	#									
Att1.2[a]	#	#									
Att2.1[a]			#	#							
Att2.2[a]			#	#							
...					#						
Att1.1[b]						#	#				
Att1.2[b]						#	#				
Att2.1[b]								#	#		
Att2.2[b]								#	#		
…										#	
Cost											

Note: Att t.l stands for the level l of attribute t, and

a: Base Version Samples

b: Text-only or Image-only Samples

#: Free parameters to be estimated.

Table A2. Differences of Mean Estimates between Treatments

		Base vs. Text-only		Base vs. Image-only	
Attribute	Level	Difference	t-ratio	Difference	t-ratio
Diversity	Medium	0.9747***	3.2338	1.3511***	3.9730
	High	1.1823***	3.3230	1.2988***	2.9245
Water	Sometimes	-0.5724***	2.7879	-0.3667*	1.8051
	Always	-1.4278***	3.8056	-1.5245***	3.7512
Mowed	0%	1.3351***	2.8891	0.6158	1.3386
	30%	0.4949	1.6149	0.6450**	2.0542
	70%	0.7402*	1.9080	1.3001***	3.4807
Planting	Medium	-0.2217	1.0808	0.1299	0.5625
Pattern	Low	0.3710	1.1010	0.4428	1.2142

*, **, ***: Significant at the 10%, 5%, and 1% levels, respectively.