The effects of presentation formats in choice experiments

Genius Murwirapachena and Johane Dikgang

ERSA working paper 803

April 2019

The effects of presentation formats in choice experiments

Genius Murwirapachena[*][†] and Johane Dikgang[‡]

October 24, 2019

Abstract

Although stated-preference surveys take various forms, the use of either text or visuals to represent attributes is uncontroversial, and they remain the commonly used formats. While prior research has investigated the impact of these commonly used formats in other disciplines, little is known about their effects on results in terms of relative importance in environmental economics literature. We conduct surveys on households' preferences for water efficient technologies in South Africa, where we compare three presentation formats, namely text, visuals, and both text and visuals. Survey data collected from 894 heads of households in the Gauteng province is analysed using the mixed logit model to test whether these three formats generate differences in estimated utilities and willingness to pay. This research sheds light on how to develop a valid presentation method for attribute levels in choice experiments, which is critical considering most environmental economics goods and services are not traded in the market. Our results show that the visuals format generates more statistically significant coefficients than the other formats. This suggests that the presentation format has significant impacts on choice. The choice between the three elicitation formats may imply a trade-off in choice precision. Our findings suggest that more research on presentation formats in environmental economics is warranted.

1 Introduction

In most choice experiment (CE) studies in environmental economics, the attributes of non-traded environmental goods are communicated to respondents in the form of a table consisting of verbal descriptions (written text). The table

[*]**Corresponding Author:** School of Economics and Development Studies, Nelson Mandela University, Port Elizabeth, South Africa. Email: murwiragenius@gmail.com

[†]Department of Public Management and Economics, Durban University of Technology, Durban, South Africa.

[‡]Public and Environmental Economics Research Centre (PEERC), School of Economics, University of Johannesburg, Johannesburg, South Africa. Email: jdikgang@uj.ac.za

normally consists of attributes, their detailed descriptions, and levels. Thereafter, respondents are presented with a series of choice sets, which are often in the form of written text. It is assumed at this stage that the respondents can fully comprehend the attributes and the attribute levels. Respondents are expected to form their preferences in response to the information provided to them in choice sets pertaining to the environmental good in question (Bateman et al., 2009).

Respondents form their preferences by cognitively combining the utilities they derive from the attribute levels that make up choice alternatives according to some function. Information plays an important role in the formation of preferences, particularly for the estimation of value for non-traded environmental goods/services, where experience of the good/service and the hypothetical market may be limited (Munro and Hanley, 2001). According to Green and Tunstall (1999), the accuracy and 'face-value' comprehension of information provided to respondents with the non-traded valuation studies should not be taken for granted.

We argue that in addition to the information presented to respondents, the presentation format in which the information is conveyed to respondents may also influence how preferences are formed. Presentation format pertains to the way the attribute alternatives are presented. This aspect is taken for granted in the environmental economics literature. Evidence on the influence of format on how preferences are formed has been observed in the literature of other fields, such as housing (see Timmermans and van Noortwijk, 1995; Wang and Li, 2004), urban planning (see Jansen et al., 2009) and consumer studies (see Townsend and Kahn, 2013). This literature attempts to address concerns about whether respondents can truly articulate their preferences, if their responses are an artefact of the experimental task, and if they can fully comprehend the typical presentation format often used to convey attribute levels, which can be complex – particularly for unfamiliar goods.

It has been argued that the presentation of attribute levels may be captured better graphically or visually. On the other hand, respondents may pay more attention to certain features of the visuals in the experiment. Moreover, some attribute levels (such as cost or other monetary attributes) may not lend themselves to visual representation (Orzechowski et al., 2005). Our study is designed to contribute to the limited but growing literature pertaining to whether presentation formats matter in choice experiments. To be specific, our study reports on the findings of a test on whether written text, visual representations or a combination of written text and visuals for attribute profiles in choice experiments generate differences in estimated empirical results.

This paper attempts to investigate whether the text, visuals or text-and-visuals presentation formats matter for discrete choice experiments in environmental and resource economics. The objective of the study is to test whether presenting attribute levels in these three presentation styles generates significantly different results with respect to attribute interpretation, relative importance, probability of adopting water conservation technologies, and willingness to pay estimates. Households completed a CE questionnaire that contained three versions of the same six choice tasks with three alternatives (status quo, alternatives 1 and 2), in which the attribute levels were presented in text, visuals or text-and-visuals. The status quo was undefined, as only the households knew their current situation. Five attributes relating to the decision of households to adopt water-saving technologies were included. Mixed logit models were used to estimate the relative importance of the attribute levels.

The contribution of this study is twofold. Firstly, more presentation formats are evaluated than in most studies in the literature; most studies compare either written text to visuals, or written text to a combination of text and visuals (see Jansen et al., 2009; Muller et al., 2010; Patterson et al., 2017; Townsend and Kahn, 2013). Our study compares three formats. Secondly, to the best of our knowledge, this is the first study in environmental economics to examine the impact of the way in which attribute profiles are presented. Evidence from other disciplines suggests that presentation formats matter (see Arentze et al, 2003; Bateman et al., 2009; Orzechowski et al., 2005). It is not clear in the environmental economics literature how to develop valid presentation methods for attribute levels in choice experiments; the investigation carried out in this study therefore sheds light on this, and so doing, contributes to the establishment of guidelines to developing valid presentation formats.

Presentation formats such as visuals improve respondents' understanding of the goods/services involved. However, considering the nature of environmental economics goods and services, coming up with the most appropriate visuals requires a lot of effort and resources. If less appropriate visuals are used to depict attribute profiles, there is a good chance that such visuals could contain various distracting effects that might bias respondents' choices (Scarpa et al., 2009). In such instances, a text presentation might be better, as it would provide clearer and more precise descriptions of the environmental good/services involved. A combination of both efforts may yield even better results, as it would combine the strengths of both approaches. Considering the effort and precision needed to come up with the most appropriate visuals, it is important to examine whether such effort does in fact improve the quality of data collected.

The rest of the paper is organised into seven sections. Section 2 reviews the literature. Section 3 presents the experimental design. Section 4 discusses the case study. Section 5 discusses the modelling approaches. Section 6 presents the experimental data. Section 7 presents and discusses the empirical findings, and section 8 concludes the study.

2 Literature on presentation formats

The issue of presentation formats is not new. Many studies in the neuropsychology literature discuss the various merits and demerits of presenting survey instruments as text, visuals, or a combination of both. Early contributions on presentation formats are found in Holbrook and Moore (1981), Childers and Houston (1984) and MacInnis and Price (1987). These studies explain how respondents process information presented either as text or visuals. A common conclusion is that information presented as text and information presented as visuals are processed differently, and by different areas of the brain. Over the years, the discussion regarding presentation formats has continued in the literature; however, only in recent years have studies emerged in the literature that examine the role of presentation formats in the choice experiment domain.

Several advantages of text presentations are discussed in the literature. The commonly identified advantage is that text presentations provide clear and appropriate descriptions of attributes and levels, as they do not have the problem of attribute interaction associated with visual presentations. Typically, text presentations do not result in the distracting effects intrinsic to visual presentations (Scarpa et al., 2009). According to Vriens et al. (1998), text presentations facilitate judgment, making it possible for respondents to make real trade-offs between given attributes and levels. These advantages are consistent with the psychological literature, which suggests that visuals dominate attributes in terms of colour and form, and may distort responses (see Holbrook and Moore, 1981; Wittink et al., 1994). As such, Patterson et al. (2017) and many other studies recommend only using visuals ahead of text when it is absolutely necessary.

However, the literature also demonstrates the advantages of visual representation. Predominantly, the use of visuals is supported in psychology literature by studies that argue that respondents are inclined to process images more readily than written text (Berlyne, 1971; Childers and Houston, 1984; Hetherington et al., 1993; Wohlwill, 1976). It is argued that visuals improve respondents' understanding and comprehension of the survey instrument. This is because it is a relatively less irritating cognitive process to perceive cues depicted in visuals than to perceive those in text (Fitzsimons et al., 2002). This assertion is consistent with the argument in Childers and Houston (1984) that advertisements presented using visual representations are remembered easily compared to those provided by verbal representations.

There is no consensus in the literature on which of the two formats to adopt when designing survey instruments. In the choice experiments literature, written text is the commonly used format for presenting choice sets (see Abdullah and Mariel, 2010; Arentze et al., 2003; Bhaduri and Kloos, 2013; Lanz and Provins, 2015; Vásquez et al., 2012). However, some choice experiment studies combine both text and visuals, to capitalise on the individual benefits of each presentation format (see Kanyoka et al., 2008; Snowball et al., 2008; Saldías et al., 2016). Importantly, most of the studies that combine text and visuals only include visuals in the table when explaining the attributes and levels. Very few include visuals in the actual choice sets. Where visuals are included in the choice sets, they are normally limited to attributes, with very few studies including them in the profiles of each choice set. Our study includes visuals to represent both the attributes and the attribute levels in choice profiles.

A clear link between presentation format and the preferences of respondents does not exist in the choice experiments literature. This has prompted an emerging interest in examining the impact of presentation formats; but the studies that test the impact of presentation formats in choice experiments are mostly in domains other than environmental economics. Most of these studies report inconsistent results (see Bateman et al., 2009; Jansen et al., 2009; Lovett et al., 2015; Muller et al., 2010; Patterson et al., 2017). While some report that text presentations and visual representations affect empirical results, others show that using either of the two formats has no meaningful effect on empirical estimations. Our study joins this debate, assessing the impact of text presentations, visual representations, and text and visual representations together on empirical estimates. By including an experiment in which information is presented only in visuals, our study is a step ahead of most similar studies in the literature.

Jansen et al. (2009) test the impact of including visuals in choice experiments on housing preferences. The study reveals that including visuals in the choice sets led to several differences in the results compared to those from text presentations. These differences are explained as emanating from accidental details in the images. Coming from a different angle, Bateman et al. (2009) test the impact of text presentations and visual representations in a choice experiment on coastal land use. The study found that text presentations generated higher gain/loss asymmetry than visual representations. Differences between results from text presentations and visual representations are further confirmed in Syrengelas (2017), in which different presentation formats yielded different welfare estimates. Other choice experiment studies that find presentation formats to affect empirical results include Muller et al. (2010) and Orzechowski et al. (2005). Most of these studies find that visual representations tend to produce more parameter estimates that are statistically significant and with larger absolute coefficients, compared to text presentations.

On the other hand, Patterson et al. (2017) use choice experiments on preferences for landscape and urban planning to test the impact of presentation formats, and find no evidence of major differences between the results from text presentations and those from visual representations. The study reports that respondents' preferences in the text survey were based on their mental images; whereas in the visuals survey, preferences were based on the displayed images. Despite this, similar results were reported from the two separate experiments. Findings from Patterson et al. (2017) are consistent with earlier work by Arentze et al. (2003), which used choice experiments on choice of transport mode to test the impact of presentation formats on empirical results. The study revealed that including visuals to text for attributes affected neither the error variance nor the measurement of attribute weights; as such, the effort it takes to develop pictorial material is not compensated for by better-quality data.

A large number of current studies in the literature attempt, in one way or the other, to examine or discuss the role of presentation formats in choice experiments. Notable studies in the literature include the work of Johnston et al. (2017) which proposes some best practices for stated preference studies, and Norman et al. (2016) which assesses clarity, difficulty, and respondent preference between presentation formats. Other recent studies examine whether packaging and presentation are key issues to consumers (see Eldesouky et al., 2016; Talati et al., 2017; van Loo et al., 2015). Notably, literature exists on choice experiment studies that seek to explain recognition heuristic, attribute attendance based on presentation formats, and the visual processing of information by respondents (see Balcombe et al., 2015; Engin and Vetschera, 2017; Krucien et al., 2017; Michalkiewicz and Erdfelder, 2016). Most of these and other similar studies report an assortment of findings. While other studies provide evidence of stability in the use of recognition heuristic, others report a mismatch between information representation and cognitive style, while others suggest a relationship between visual attention and individuals' preferences which depends on the type of product attribute.

It is evident from this section that the existing choice experiment literature on presentation formats is predominantly in domains other than environmental economics. Our current study attempts to address this by examining concerns around the way choice experiment profiles are presented in environmental economics. As revealed in this section, the few emerging choice experiment studies in the literature compare text-only experiments to those employing text and visuals. Our study is a step ahead of these studies because in addition to comparing text experiments to text-and-visual experiments, it also compares a visuals-only experiment to the other two. More precisely, our current study compares three presentation formats, as opposed to the more common procedure of comparing only two formats. In comparing three presentation styles, hopefully the study will shed light on the extent of the bias found in the empirical estimates produced by the environmental economics literature.

3 Case study: households' willingness to adopt water-saving technologies

One of the biggest criticisms of environmental and resource economics choice experiments is that the goods and services being evaluated are not traded in the market, so respondents are not familiar with them; hence they may find making trade-offs very difficult. This suggests that the behaviour underlying CE results is not well understood. It is therefore possible that by default, respondents may resort to a simplified decision rule – particularly in instances where choices are too difficult. There are ongoing debates about the complexity of choice tasks in environmental-related choice experiments, and the extent to which respondents can comprehend choice tasks as intended in the experiment. Despite this, there is increasing use of CE experiments in environmental economics. Since some of these experiments may be used for policymaking, the accuracy and validity of the measured preferences are key to avoiding incorrect policy choices based on invalid experiments.

Because of global warming and growing water scarcity, policymakers are increasingly exploring ways to conserve water. As households are among the biggest water users, they are often targeted by decision-makers using a variety of tools, some of which are intended to change user consumption behaviour through the adoption of conservation technologies. The case study in this study therefore has serious policy implications; hence, it is essential that participants in the experiment understand the choice tasks as fully as possible, to reveal their true preferences. It is critical that participants in CEs better understand the included attribute-related information, in order to make choices that reflect preferences more accurately.

A failure on the part of the households in our experiment to understand the adoption of water-saving technologies will put the experiment at risk. In this area, little is known about how much of the attribute-related information respondents understand. This is worrisome, considering the increasing use of this technique in environmental economics. One of the reasons for failure to fully comprehend the experiment is lack of understanding of the attribute levels in the experiment. There is great diversity in the way environmental economics information is translated into attribute levels, in how they are explained to participants, and in how choice tasks are presented in CEs on environmental or environmentally related topics.

It is argued that the presentation format used in an experiment may impact the respondents' understanding of the information on attribute level contained in the experiment. Considering that in environmental economics most studies use text only, it is plausible that visual attribute-level communication might help to make choice tasks easier, and therefore improve the quality of respondent feedback. It is also plausible that attribute-level information that contains both text and visuals may yield more consistent results. The question is whether any presentation format gives more consistent responses, and whether the participants would prefer one format over another. There is no empirical evidence on this score in the environmental economics literature. The adoption of water-conservation technologies can be presented easily using either format, so it is deemed suitable for the purposes of this study. We therefore test for differences in CE results using the three presentation styles.

4 Experimental design

The first step in choice experiments involves selecting relevant attributes and assigning realistic levels to each attribute. Selected levels assigned to each attribute should be feasible, realistic, non-linearly spaced, and should span the range of respondents' preference maps (Hanley et al., 2001). Both attributes and levels can be deduced from a literature review, focus groups, pilot studies, and expert consultations. After attributes are identified and relevant levels are assigned to each attribute, experimental design commences. Experimental design is explained as the specialised and scientific manipulation of the levels of one or more attributes to generate choice profiles (Hensher et al., 2015). The most common classes of experimental design in the literature are full factorial, orthogonal, and efficient designs. This section discusses the attributes and levels used in the study, as well as how they are experimentally designed into choice profiles.

4.1 *Attributes and levels*

In determining the attributes for our experiment, we used a combination of both literature review and expert consultation. The literature shows that household water-efficient technologies can be categorised based on areas in a home. A typical South African middle-income household of four spends 25% of their water use in flushing the toilet, 25% on garden and outdoor activities, 24% on bathing or showering, 13% on laundry, 11% in the kitchen, and 2% on other activities (Price, 2009). We use these areas as attributes and adopt the various technologies that may be fitted into each of these areas as levels. From a series of expert consultations, four key areas were adopted as attributes, namely kitchen, shower, toilet and garden/outdoor. We include the monthly water bill as our monetary attribute essential for measuring social welfare.

Several water-efficient technologies that can be installed to save water in a homestead are identified in the literature (see Hering and Ingold, 2012; Jones and Hunt, 2010; Makki et al., 2013; Mini et al., 2015; Still and Bhagwan, 2008; Willis et al., 2013). After consulting with experts, our study has adopted water-efficient technologies that are deemed necessary in the South African context. For the kitchen devices, as levels we use efficient dishwashers, efficient taps, and a system for collecting used water. In terms of shower devices, we use efficient showerheads and shower timers as levels, while for toilet devices we use dual-flush cisterns, interruptible (multi) flush cisterns, and cistern displacement devices (hippo bags) as levels. For garden/outdoor devices, the levels are time-based irrigation controllers, micro-drip irrigation systems, and water tanks for harvesting rainwater[1].

Investing in water-efficient technologies essentially reduces a household's monthly water bill. Therefore, the various possibilities for reduced monthly water bills are used as levels for the monetary attribute. To determine these possibilities, we consider the average monthly water bill for households in the study area. Using data collected by the National Income Dynamics Study (NIDS), an average of R450 (around \$31.47)[2] per month was determined to be the current average water bill. If households were to adopt water-saving technology, their bill would be reduced by 75%, 50% or 30%, that is, from R450 to one of R110, R225 or R315 (\$7.69, \$15.73 or \$22) per month, respectively. The final list of attributes and levels is presented in Table 1.

The attributes and levels presented in Table 1 are experimentally designed into choice set profiles. In addition to the designed profiles, we also include a status quo (SQ) profile. An SQ profile essentially avoids the undesirable effects associated with forced choices (Dhar and Simonson, 2003; Ferrini and Scarpa, 2007). Our study uses an undefined SQ in each choice set. An undefined SQ is an individual-specific SQ in which where each respondent envisages their

[1] We agree that the technologies chosen as levels may also be used as attributes in other studies. However, in the context of our study the emphasis is on the areas in a home where households can save water by installing efficient technologies (i.e. kitchen, shower, toilet and outdoor). As such, water-efficient devices that can be fitted in these areas are used as levels in our choice experiment.

[2] As of the 24^{th} of October 2018, US\$1 = ZAR14.30.

own current status and compares it to the experimentally designed hypothetical options (Hess and Rose, 2009). Undefined SQs and similar approaches are common in the literature (see Campbell et al., 2008; Hess and Rose, 2009; Marsh et al., 2011; Scarpa et al., 2007; Train and Wilson, 2008; Willis et al., 2005). They are commonly used when it is difficult to ascertain the current situation for the sample. This is the case in our experiment, because we cannot determine the current use of water-efficient technologies with certainty. In South Africa, household water-conservation practices are not clearly documented. The use of an individual-specific SQ essentially avoids the problems associated with the risk of imposing an inapplicable SQ, which could have been the case in our study.

4.2 *Description of design*

This study uses an efficient design to generate the choice profiles presented to respondents. Efficient designs are praised in the literature for producing robust data that give more reliable parameter estimates with an even lower sample size than designs such as full factorial and orthogonal. According to Rose and Bliemer (2009), efficient designs give smaller widths of confidence intervals observed around the parameter estimates, and maximised asymptotic t-ratios for each parameter, thereby improving the reliability results. However, efficient designs are only efficient if prior parameters are known. If incorrect prior parameters are used, efficient designs become inefficient (Bliemer et al., 2008). To address this problem, the literature recommends drawing parameter estimates using the Bayesian parameter distributions. Bayesian parameter estimates are sensitive to misspecification of priors, because they assume prior parameter values to be approximately known and randomly distributed. Using a D-*error* efficient measure with prior parameters drawn by means of Bayesian parameter distribution, the design becomes a Bayesian D-*error* design (i.e. D_b-*efficient*). This Bayesian D-*error* design is commonly used to examine efficient designs where the true population parameters are not known with certainty.

Using a normally distributed Bayesian D-*efficiency* design, this study experimentally designs six choice sets of two profiles each. Following suggestions by Bliemer et al. (2008) that the Gaussian method is the best approximation method for Bayesian efficient designs, we adopt the Gaussian method to come up with the number of draws for Bayesian priors. The rule of thumb for determining the absolute minimum Gaussian quadrature is 2^K, where K is the number of Bayesian priors. Given the number of attributes and levels in our experiment, we use the maximum possible Gaussian draws (i.e. 32 draws). These Gaussian draws are used in the normally distributed Bayesian D-efficient design adopted to populate the six choice sets of two profiles each that are experimentally designed in this study. The generated choice profiles are then presented to respondents using the three presentation formats being tested in this study. An example of a choice set used in the text experiment is given in Table 2.

Respondents were asked to select their preferred profile from options 1 and 2, or they could opt out by choosing the status quo option. The same was done in the second experiment, in which attributes and levels were presented as visuals (except for the levels of the monetary attribute). Studies that use only visuals are rare in the literature, where most studies add text to visuals. The scarcity of such studies could be attributed to the disadvantages of visuals that are highlighted in the literature (see Holbrook and Moore, 1981; Vriens et al., 1998; Wittink et al., 1994). However, there are also several advantages of visuals over text representations, making it imperative to compare empirical results derived from the two formats. An example of the visual presentation choice set used in this study is given in Table 3.

The visuals presented in each profile of Table 3 represents the same information that was presented as text in Table 2. Additionally, the designed choice profiles were also presented as both text and visuals. Our text-and-visuals experiment is a step ahead of most choice experiment studies, which only include images in the attributes (see Kanyoka et al., 2008; Snowball et al., 2008). Our choice sets include images in the profiles. Saldías et al. (2016) used this style to present choice sets in a study that elicited farmers' preferences for wastewater re-use frameworks in agricultural irrigation. Table 4 gives an example of the choice sets used in the text-and-visuals experiment.

In addition to the choice experiment, our questionnaire also included other sections. The second section collected general information on households' water conservation behaviour and technology. Such information is essential, because a relationship between water-efficient technology and water-use behaviour is identified in the literature (see Davis, 2008; Freire-Gonzalez, 2011; Ghosh and Blackhurst, 2014; Smeets et al., 2014). The literature argues that households adopt non-efficient habits when they install efficient technologies. The third section collects the socio-economic characteristics of the respondents, which are essential when establishing the drivers of respondents' choices. Various socio-economic characteristics of respondents are identified as key determinants of choices in the literature (see Millock and Nauges, 2010; Martinez-Espineira and García-Valiñas, 2013; Pérez-Urdiales and García-Valiñas, 2016)[3].

5 Modelling

Developed from the random utility theory, choice experiments assume that individuals are rational decision-makers who choose the most preferred (utility-maximising) option when faced with a possible set of options (Abelson and Levy, 1985; Howard, 1977; McFadden, 1973). According to McFadden (1973), these rational individuals make choices based on the characteristics of the good, along with a random component. The random component could emerge from the uniqueness in the individual's preferences, or due to researchers having incomplete information about the individual observed (Ben-Akiva and Lerman, 1985). The literature proposes that the utility derived from an option by an

[3]It is important to note that the three questionnaires used in this study only differed in the formats used to present the choice experiment section. The information presented was similar across all three questionnaires.

individual is not known, but can be decomposed into a deterministic component and an unobserved random component, as follows:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \tag{1}$$

Parameter U_{ij} represents the utility of individual i obtained from option j, parameter V_{ij} is the deterministic component which is normally specified as a linear index of the attributes in a choice set, and ε_{ij} is the unobserved random component of latent utility which captures the consequence for choice of uncertainty due to incomplete information. Equation 1 represents the basic utility function and may be expressed by decomposing the indirect utility function for individual U_{ij} into two main components. If that occurs, the utility function then becomes:

$$U_{ij} = V_{ij}\left(X_{ij},\, C_{ij},\, \beta\right) + \varepsilon_{ij} \tag{2}$$

Equation 2 decomposes V_{ij} into attributes X_{ij} and C_{ij}, where X_{ij} is the vector of non-monetary attributes associated with option j, while C_{ij} is the monetary attribute of option j, β is the vector of preference parameters for the population in the sample, and ε_{ij} is the stochastic component. Random utility posits that any rational individual i chooses option j over option k if $U_{ij} > U_{ik}$. Each option consists of a bundle of attributes. When an individual selects one option over the other, it suggests that the hypothetical utility derived by the individual from the chosen option is greater than the utility of the other option not chosen (Greene, 2003; Louviere, 2001). Therefore, the probability P_i of selecting option j because $U_{ij} > U_{ik}$ is illustrated as:

$$P_i\left(j\right) = \text{Prob}\left(V_{ij} + \varepsilon_{ij} > \ V_{ik} + \varepsilon_{ik}\right) \ \ \forall \ \ k?C,\, k \neq j \tag{3}$$

If the error terms are independently and identically distributed (IID) with an extreme value type I distribution, the variance of which is $var\left(\varepsilon\right) = \pi^2\tau^2/5$, where τ is a scale parameter that is used to normalise the model, then the choice probability of an option is expressed as:

$$P_{ij} = \exp\left(\frac{v_{ij}}{\tau}\right) / \sum_{k=1}^{K} \exp\left(\frac{v_{ik}}{\tau}\right) \tag{4}$$

Several logistic models are then used to estimate the probability defined in equation 4. The most basic of these logistic models is the conditional logit model (CLM). Also known as the multinomial logit (MNL) model if there are no choice varying attributes, the model uses the maximum likelihood estimation approach (Hensher et al., 2015). The model has enjoyed extensive use in the literature and Hensher et al. (2015) identifies it as the “workhorse” for discrete choice experiments. However, MNL is criticised for assuming respondents to have homogenous tastes for observed attributes, and that the random part of utility obeys the independence from irrelevant alternatives (IIA) as well as the independence and identical distribution (IID) properties. These assumptions are unrealistic as they rule out persistent heterogeneity in tastes for observed and unobserved product attributes (see Greene, 2012, Hensher et al., 2015, Keane and Wasi, 2012).

Models that address the criticisms of MNL include the mixed logit (MXL – also known as the Random Parameter Logit - RPL) model. This study uses the MXL model to examine the impact of presentation formats on empirical results. MXL is also used in many other studies that examine presentation formats (see Caussade et al., 2005; Syrengelas, 2017; Patterson et al., 2017). It allows coefficients to vary randomly across individuals, reflecting the reality that different respondents have different tastes and preferences for attributes in each choice set. The many advantages of MXL include its ability to account for both observed and unobserved heterogeneity in the preference parameters; and that it is versatile, with both single cross-sectional and panel data (Hensher et al., 2015). MXL breaks down coefficients into a population mean, and an unobserved individual's deviation from that mean (Greene, 2012), as follows:

$$U_{ij} = \beta X_{ij} + \eta_{ij} + \varepsilon_{ij} \tag{5}$$

where β is the population mean, while η_{ij} is the individual deviation from the population mean which shows the individual specific heterogeneity, with mean zero and standard deviation one (Greene, 2012).

As is common in the literature, we also test for variations in welfare measures by examining the marginal willingness to pay (MWTP) estimates across the three presentation formats. MWTP estimates show the marginal rate of substitution (MRS) between each attribute and the monetary attribute; this is an important output of choice models, as it gives average estimates of what respondents are prepared to pay for or against each attribute (Hensher et al., 2015). The MWTP for attribute X is calculated by taking the ratio of the derivatives of attribute X and the monetary attribute C, which in the case of a linear function is given as:

$$WTP_X = \frac{\Delta X}{\Delta C} = -\frac{\frac{\partial U_{ij}}{\partial X_j}}{\frac{\partial U_{ij}}{\partial C_i}} = -\frac{\beta_j}{\mu} = MWTP \tag{6}$$

The MWTP presented in equation 6 is a simple ratio of the coefficients of the parameter estimates, which can be compared across models because the scale parameters are cancelled (Hensher et al., 2015). In the context of our study, it is essential to assess the impact of the presentation formats on welfare measures. This is because the literature reports inconsistent results on variations in the actual preferences and welfare measures (see Fitzsimons et al., 2002; Jansen et al., 2009; Lovett et al., 2015; Orzechowski et al., 2005; Patterson et al., 2017; Ro et al., 2009; Scarpa et al., 2009; Syrengelas, 2017; Vriens et al., 1998). Therefore, we examine whether variations in the MWTP would be observed across the three different presentation formats.

6 Experimental data

6.1 *Data collection*

The study is based on experimental data collected from 894 heads of households in the Gauteng Province, during the period November to mid-December 2017 and mid-January to February 2018. Survey instruments were prepared in English, and enumerators conversant in both English and other local languages were recruited from residents in the study area. These enumerators were trained and supervised during the data collection process. The survey, the main aim of which was to elicit the impact of presentation formats, collected stated-preference data on household preferences for water-efficient technologies. A split-sample survey was adopted, in which the first sub-sample was presented with a text experiment and data was collected from 232 respondents. The second sub-sample was presented with the same information, but using visual representations, and 257 complete responses were collected. The third sub-sample was presented with a questionnaire that combined both text and visual representations, and 405 complete responses were collected.

Respondents in each sub-sample were from the same residential area, hence their socio-economic characteristics could be expected to be similar. This is typical of South Africa, where historically, residential areas are clustered, mainly for socio-economic and historical reasons. Enumerators spent a week in a residential area collecting data using the first survey instrument. Once certain expected data points were achieved, enumerators would then move to another area, still using the first instrument. After collecting enough data from our targeted residential areas using the first instrument, enumerators went back to the same areas with the second questionnaire. However, on the second visit, different households from those interviewed in the first survey were interviewed. This process was followed until enough data points had been collected using the three questionnaires.

6.2 *Descriptive statistics*

The questionnaires used to collect information were made up of three sections. In the first section, respondents were presented with the choice experiment. The second section collected some general information on households' current water-use behaviour, as well as their current use of water-efficient technology. Such information is essential in determining people's choices. For example, households without water-efficient technologies installed – hypothetically – would prefer changes to their current water appliances compared to those with efficient technologies currently installed in their homes. The third section collected the biographical information of the respondent. The biographic characteristics collected include the respondent's gender, household size, education, age, marital status, race, income and source of income. The literature identifies these variables as key drivers in how respondents process information.

The information collected in the second and third sections of the questionnaire is essential in our study, which uses a split-sample approach and compares empirical results across sub-samples. For us to be able to compare empirical results across sub-samples, there should be some similarity and consistency in the socio-economic characteristics of the respondents across the sub-samples. The descriptive statistics of these biographical characteristics are given in Table 5 below. In addition to the three sub-samples, the table also presents statistics on pooled data[4].

Except for the number of respondents, where the text-and-visuals experiment had more respondents than the other two experiments, the descriptive statistics presented in Table 5 above show some consistency in socio-economic characteristics across the three experiments. Across all the experiments, there were slightly more male respondents than female respondents. Equally, in all the experiments there were more respondents belonging to the 'black' racial group than to the other groups. Most of the respondents in all three experiments had high school education and receive salaries or wages as their main source of income. The consistency of the socio-economic characteristics across the three experiments makes it possible for us to compare the empirical results estimated in each experiment.

6.3 *Frequency distribution of efficient technologies and water-consumption habits*

The current use of water-efficient technologies by households may have an impact on their choices. In South Africa, little is known about households' water-consumption behaviour, and the extent to which they make use of efficient technologies. Therefore, eliciting such information is essential for creating new knowledge useful for policymaking. Equally, the literature suggests a link between water-use behaviour and the installation of water-efficient technology. We asked respondents to indicate whether they currently have water-efficient technologies installed. Eight questions on water-efficient technologies were asked, using a four-point Likert scale with the options 'Yes', 'No', 'Not applicable' and 'Not sure'. To be specific, respondents were asked to indicate whether they currently have water-collection tanks, cistern displacement devices, water-flow regulators, efficient showerheads, efficient toilet cisterns, multi-flush toilet cisterns, dishwashers, and/or efficient garden devices. Except for the efficient toilet devices, the modal response for all technologies was 'No', indicating that households in our sample did not currently use water-efficient technologies. This result was observed consistently across all three experiments.

Furthermore, we elicited the possible reasons for not installing efficient technologies. Although there could be various reasons, respondents were asked to choose between 'I cannot afford', 'I did not know about them', 'I have no infrastructure to connect them', 'They are not important to me', and 'Other'. We assume that the reason for not installing water-efficient devices has an impact

[4] Pooled data is a combination of data from the three presentation formats. It is possible to combine these datasets because the information in all the three questionnaires is similar, i.e. the three questionnaires collected the same information.

on both the respondents' choices and the format used to present the attribute levels. For example, respondents who 'did not know about water-efficient devices' are more likely to make informed decisions when the technologies are presented both textually and visually because presenting the technology as text only may not give enough information. Equally, respondents who 'do not have the infrastructure' to install certain technologies are likely to ignore choice profiles that contain such technologies. Therefore, it is imperative to elicit such information. Figure 1 presents the frequency distribution of the reasons for respondents' not installing efficient water technologies. The frequency distribution is reported for each of the three experiments.

Across all three experiments, the main reason for not adopting water-efficient technologies is that households 'cannot afford the technology'. Interestingly, respondents also indicated that they 'did not know about water-efficient technologies'. This justifies the assertion by Vloerbergh et al. (2007) that the nature of water makes it a low-involvement product, such that people do not think about it as long as it is available and does not have colour or smell or taste odd. Figure 1 above shows that the main reasons for not currently installing water-efficient technologies are consistent across the three experiments. This clearly shows some similarity in the respondents sampled in each experiment, which makes it possible to compare estimation results across the experiments.

Additionally, respondents were asked eleven behavioural questions, using a four-point Likert scale with the options 'Never', 'Once in a while', 'Always' and 'Not applicable'. The eleven questions asked were on inefficient water-use behaviour; if respondents indicated 'Never', it showed that they were practising efficient water-use behaviour, while if they indicated 'Always', it showed they were practising inefficient water-use behaviour. Generally, we observed that households in the sample practised efficient water-use behaviour. Responses were mostly consistent across the three experiments, which indicates that our sampled respondents possessed almost similar characteristics across sub-samples.

Finally, we present the frequency distribution of the stated preference choices. A presentation of the frequency distribution of how each alternative was chosen in each experiment is important when checking if choices were consistent across presentation formats. Where consistent choices are observed across formats, it makes the comparison of empirical estimates possible. However, if inconsistences are observed across formats, it implies that the presentation format affected respondents' choices. This information is shown in Figure 2 below.

Figure 2 above shows that options 1 and 2 had an almost equal chance of being selected by respondents, implying that there were real trade-offs between the two options. In most choice experiments, the problem of status quo bias is reported, where respondents resort to choosing the status quo option they know as opposed to hypothetically designed options (Lanz and Provins, 2015; Marsh et al., 2011; Samuelson and Zeckhauser, 1988). The problem of status quo bias makes econometric analyses complex, and is alleged to bias estimation results (Meyerhoff and Liebe, 2009). In our current study, we avoided this problem by using an individual specific status quo (see Campbell et al., 2008; Hess and Rose, 2009; Marsh et al., 2011). The distribution of choices shows real trade-offs between options 1 and 2, which will make results from our econometric analyses more robust. Importantly, the distribution of choices is consistent across the three experiments, which makes it possible for us to compare empirical estimates across experiments. The next section presents and discusses the estimation results based on the stated preference data.

7 Empirical findings and discussion

To examine the impact of presentation formats on empirical estimates, we use the mixed logit (MXL) model as an estimation tool. Since our study conducted three experiments, we estimate utility functions for each of the three experiments. The MWTP estimates are also estimated for each experiment. The rationale of these two important analyses is to compare estimates across experiments and see if there are variations in terms of the statistical significance, sign and magnitude of the estimates. This section presents and discusses the estimation results. To estimate utility functions, the study adopts unconstrained MXL models where the five attributes of the study are modelled as normally distributed random parameters while alternative specific constants (ASCs) are modelled as fixed parameters. Results are obtained using the Halton sequence for simulation, based on 1000 draws. Utility models estimated in this study are defined as:

$$\begin{aligned} U_{ij} = \beta_0 + \beta_1 KITCHEN_{ij} + \beta_2 SHOWER_{ij} + \beta_3 TOILET_{ij} + \\ \beta_4 GARDEN_{ij} + \beta_5 BILL_{ij} + \varepsilon_{ij} \end{aligned} \tag{7}$$

Parameter β_0 represents the ASCs, while parameters β_1 to β_5 are coefficients of attributes and ε_{ij} is the random error component. The utility function presented in Equation 7 is estimated for the three experiments. The estimation results are presented in Table 6.

We interpret the estimation results presented in Table 6 above based on the sign, magnitude and statistical significance of the random parameters. The parameter estimate of each attribute indicates the utility derived by respondents. To be specific, the sign of the parameter estimate shows the direction of the relationship between an attribute and the respondents' utility derived, while the magnitude of the parameter estimate shows the extent of the impact. The statistical significance of the parameter estimate shows the importance of an attribute to respondents.

Positive parameter estimates show that respondents prefer improvements in the attribute, whereas negative estimates show that respondents do not prefer improvements. Using the attribute parameter estimates reported for KITCHEN devices in all three models, the results are interpreted to mean that households prefer improvements in the KITCHEN devices. In the text model, for example, a unit improvement in the KITCHEN devices will increase respondents' utility by approximately 0.01 units, that is, a 10% improvement in KITCHEN devices increases the respondents' utility by about 0.1%. Regarding the negative attribute parameters, a unit increase in the BILL, for example, reduces the respondents' utility by approximately 0.01 across all the three models. This implies that when making choices, respondents did not prefer alternatives with higher water bills.

Variations in the sign and magnitude of parameter estimates across experiments are interpreted to mean that presentation formats affect empirical results. Results show negligible differences in the magnitude of the parameter estimates for each attribute across the three experiments. The magnitudes of the parameter estimates are well within the same range, in absolute terms. This result is consistent with findings from Arentze et al. (2003) and Patterson et al. (2017), where the size of the coefficients in absolute terms showed little difference across different experiments. However, our estimates are not consistent with results in similar studies that show large coefficients for visuals experiments compared to text experiments (see Bateman et al., 2009; Orzechowski et al., 2005; Vriens et al., 1998).

Our empirical results show considerable differences in the signs of attribute parameter estimates across experiments. Only KITCHEN and BILL reported parameters with the same signs across the three experiments. Nevertheless, some similarities are observed when comparisons are made between any two of the three experiments. For example, SHOWER has the same sign in the visuals and the text-and-visuals experiments, TOILET has the same sign in the text and the text-and-visuals experiments, and GARDEN has the same sign in the text and the visuals experiments. However, if the signs of the attribute parameters are compared across all three experiments, there are noticeable differences in most of the parameter estimates. This is in line with the results in some studies in the literature, which show that presenting information as visuals is likely to give different results compared to scenarios where the same information is presented as text (see Molin, 2011; Rizzie et al., 2012; Wittink et al., 1994). The argument usually put forward is that visuals present greater evaluability, which reduces the respondents' judgement error.

An analysis of the significance of the parameter estimates presented in Table 6 above shows only two attributes that are statistically significant in the text experiment, while four attributes are statistically significant in the visuals experiment and three attributes are statistically significant in the text-and-visuals experiment. BILL is the only attribute that is statistically significant across all three experiments. KITCHEN is statistically significant in the visuals experiment and the text-and-visuals experiment but is statistically insignificant in the text experiment. SHOWER is statistically significant in the text and the text-and-visuals experiments. Except for the parameter estimates for BILL, there are no other consistent estimates between the text and the visuals experiments. These results are consistent with findings in the literature that visuals always have more statistically significant coefficients than the other presentation formats. Perfect examples of studies whose findings are consistent with ours include Jansen et al. (2009), Orzechowski et al. (2005) and Vriens et al. (1998). According to Patterson et al. (2017), most studies in the literature show visually presented variables taking on more importance than variables presented through text.

While the sign, significance and magnitude of random parameter estimates are essential when comparing empirical results, random parameter estimates themselves show the population mean. Therefore, it is also important to compare the dispersion that exists around the sample population in each format. This information is given by the standard deviations of the parameter distributions. Insignificant parameter estimates for derived standard deviations indicate that the dispersion around the mean is statistically equal to zero, suggesting that all information in the distribution is captured within the mean (Hensher et al., 2015). On the other hand, statistically significant parameter estimates for derived standard deviations of a random parameter suggest the existence of heterogeneity in the parameter estimates over the sampled population around the mean parameter estimate. According to Hensher et al. (2015), this implies that different individuals possess individual-specific parameter estimates that may be different from the sample population mean parameter estimate.

In terms of the standard deviations of random parameters, our results show that the text-and-visuals model had more estimates that were statistically insignificant than the other two models. Only two attribute parameters in the text-and-visuals model (KITCHEN and SHOWER) had statistically significant standard deviations. In the visuals model, all estimates except for GARDEN were statistically significant; while in the text model, all estimates were statistically significant. This suggests that in the text and the visuals models, different respondents possessed individual-specific parameter estimates that may be different from the sample population mean parameter estimate. However, in the text-and-visuals model the dispersion around the mean of most estimates is statistically equal to zero, suggesting that all information in the distribution is captured within the mean. This implies that the text-and-visuals experiment was able to capture the true preferences of respondents better than the other experiments.

It is also common practice in the literature to compare empirical estimates on the measures of welfare across presentation formats (see Bateman et al., 2009; Patterson et al., 2017). This section presents MWTP estimates, which are commonly used as welfare measures in the literature. MWTP estimates show the average estimates of what respondents are prepared to pay for or against improvements in each attribute. Positive and significant figures show the average amount that households are willing to pay for improvements in the attribute, whereas negative and significant figures show how much households are willing to accept as compensation for changes in the attribute. Empirical estimates for MWTP for the study are presented in Table 7.

MWTP estimates presented in Table 7 are interpreted to mean that in the text experiment, respondents are willing to pay \$2.84 for improvements in SHOWER devices. In the visuals experiment, respondents are willing to pay \$0.90, \$1.22 and \$0.89 for improvements in KITCHEN, TOILET and GARDEN devices respectively. In the text-and-visuals experiment, respondents are willing to pay \$15.59 for improvements in SHOWER devices. Two main observations are made from a comparison of the statistical significance of the MWTP estimates. Firstly, the visuals experiment has more MWTP estimates that are statistically significant than the other two experiments, which have one statistically significant MWTP estimate each. This observation is consistent with earlier results on utility functions, where the visuals experiment also emerged as having more attribute parameter estimates that were statistically significant than the other experiments.

Secondly, we observe that the MWTP estimates reported in the text-and-visuals experiment are larger in absolute terms than those from both the text and the visuals experiments. When the sizes of MWTP estimates for the text and the visuals experiments are compared, it can be observed that the latter has more estimates that are bigger than the former in absolute terms. This agrees with findings in the literature that images tend to produce estimates that are mostly bigger than those from text experiments, in absolute terms (Bateman et al., 2009; Orzechowski et al., 2005; Syrengelas, 2017; Vriens et al., 1998). Overall, we observe that MWTP estimates were largely different in terms of sign, significance and magnitude across the three presentation formats. Based on these results, we argue that the presentation format also affects MWTP estimates.

8 Conclusion

This paper uses choice experiments to examine the impact of presentation formats on empirical results. The focus of the paper was to establish whether utility functions and MWTP estimates are affected by the format used to present choice experiments. To achieve this, we used data from experiments on household preferences for water-efficient technologies in the Gauteng province of South Africa. The study compares three experiments, namely a text experiment, a visuals experiment and a text-and-visuals experiment. In the text experiment, respondents answered choice questions with alternatives presented as text, while in the visuals experiment, the same information was presented to respondents in the form of images. The third experiment provided the same information again but using both text and visual representations. By presenting an experiment that is entirely visual, our study is a step ahead of many similar studies in the choice experiment literature, which mainly compare text presentations with text and visual presentations (see Jansen et al., 2009; Orzechowski et al., 2005; Patterson et al., 2017). Our study uses the MXL model for empirical estimation, and four main findings can be reported.

Firstly, we found that while only two attributes emerged as important in the text experiment, four attributes were important in the visuals experiment and three were important in the text-and-visuals experiment. The literature explains the importance of attribute parameters as based on the size and statistical significance of the coefficients. Although there was not much difference in the size of the coefficients (in absolute terms) across the three experiments, we found that the visuals experiment had more statistically significant coefficients than both the text and the text-and-visuals experiment. This result is consistent with those of similar studies in the literature, which also find visual experiments to have more statistically significant coefficients than text experiments (see Jansen et al., 2009; Orzechowski et al., 2005; Vriens et al., 1998). Since the visuals and the text-and-visuals experiments both had more coefficients that were statistically significant than the text experiment, we argue that including visuals in the choice profiles increased the number of attributes that were important to respondents.

Secondly, a comparison of attribute parameters across all three experiments showed some differences in the signs of each parameter, with only two attributes having the same sign across all three experiments. However, a few similarities in the signs were observed when comparisons were made between any two of the three experiments. Prior to the tests, we hypothesised that although the magnitude and statistical significance of each attribute parameter may differ across experiments, the sign of each parameter should be the same. This is because descriptive statistics in our study showed similarities in the socio-economic characteristics of respondents across the experiments. Our hypothesis was shaped by Patterson et al. (2017), where no meaningful differences were observed in results across experiments. However, the results in this study confirm reports in the literature that visuals and text experiments give different results (see Molin, 2011; Rizzie et al., 2012; Wittink et al., 1994).

Thirdly, we observed that the text-and-visuals experiment reported fewer attribute parameters with dispersion around the sample population than the text and the visuals experiments. Only two parameters in the text-and-visuals experiment had statistically significant standard deviations. This indicates that the random parameter estimates reported in the text-and-visuals experiment correctly reflect respondents' choices, except for two attributes. In the visuals experiment, only one parameter was statistically insignificant; while in the text model, all estimates were statistically significant. This suggests that in the text and the visuals experiments, different respondents possessed individual-specific parameter estimates that may be different from the sample population mean. Considering this, we argue that the text-and-visuals experiment was able to capture the true preferences of respondents better than the other experiments.

Finally, the MWTP estimates showed that households were willing to pay for more attributes in the visuals experiment than in the other two experiments. In the visuals experiment, respondents were willing to pay for three attributes, whereas they were only willing to pay for one in the text experiment and one in the text-and-visuals experiment. Again, this confirms reports in the literature that visual experiments tend to have more significant parameters than text experiments (see Bateman et al., 2009; Orzechowski et al., 2005; Syrengelas, 2017; Vriens et al., 1998). A comparison of the magnitude of the MWTP estimates across the three experiments showed that the text-and-visuals experiment had larger estimates than the other two experiments, in absolute terms. Overall, the MWTP estimates were largely different in terms of sign, significance and magnitude across the three presentation formats.

Based on the results presented in this study, we join other studies in the literature in arguing that visually-presented attributes tend to take on more importance than attributes presented through text. However, we advise caution when presenting experiments as visuals, since other less important aspects such as colour and form may distort preferences. On the other hand, the text-and-visuals experiment showed some consistency with both the text and the visuals experiments in terms of the sign and significance of parameters. By combining both text and visuals, the experiment was able to clarify attributes to respondents, thereby yielding more robust stated preference data and empirical estimates. Overall, we argue that the format of presenting information matters in choice experiments conducted in environmental economics.

The main limitation of our study was that each of the three instruments used to collect information was presented to different respondents with similar socio-economic characteristics. We did not ask each respondent to answer all three experiments because that would have resulted in learning, fatigue and boredom. However, if these could be controlled, results from studies where all three questionnaires are answered by the same respondents would be interesting. Therefore, we recommend that future studies should test this by presenting all three experiments to the same respondents. Additionally, more research is required on the effects of the use of various presentation styles in environmental economics, so that guidelines can be established on how to develop valid presentation formats for attribute levels in the choice tasks found in choice experiments. It is important to appreciate that the effect of the presentation format may dependent on the context under consideration. Sometimes images can influence people's decisions more in one field than the other. Therefore, more research on this subject is warranted in the context of environmental economics.

References

[1] Abdullah, S. and Mariel, P. 2010. Choice experiment study on the willingness to pay to improve electricity services. *Energy Policy,* 38**,** 4570-4581.

[2] Abelson, R. P. and Levy, A. 1985. *Decision Making and Decision Theory. (In Lindzey, G. and Aronson, E. Eds). Handbook of Social Psychology 1.*, New York: Random House.

[3] Arentze, T., Borgers, A., Timmermans, H. and Delmistro, R. 2003. Transport stated choice responses: effects of task complexity, presentation format and literacy. *Transportation Research Part E: Logistics and Transportation Review,* 39**,** 229-244.

[4] Balcombe, K., Fraser, I. and McSorley, E., 2015. Visual attention and attribute attendance in multi-attribute choice experiments. *Journal o Applied Econometrics*, 30**,** 447-467.

[5] Bateman, I. J., Day, B. H., Jones, A. P. and Jude, S. 2009. Reducing gain–loss asymmetry: a virtual reality choice experiment valuing land use change. *Journal of Environmental Economics and Management,* 58**,** 106-118.

[6] Ben-Akiva, M. E. and Lerman, S. R. 1985. *Discrete choice analysis: theory and application to travel demand*, MIT press.

[7] Berlyne, D. E. 1971. *Aesthetics and psychobiology*, JSTOR.

[8] Bhaduri, A. and Kloos, J. 2013. Getting the water prices right using an incentive-based approach: an application of a choice experiment in khorezm, uzbekistan. *The European Journal of Development Research,* 25**,** 680-694.

[9] Bliemer, M. C., Rose, J. M. and Hess, S. 2008. Approximation of Bayesian efficiency in experimental choice designs. *Journal of Choice Modelling,* 1**,** 98-126.

[10] Campbell, D., Hutchinson, W. G. and Scarpa, R. 2008. Incorporating discontinuous preferences into the analysis of discrete choice experiments. *Environmental and Resource Economics,* 41**,** 401-417.

[11] Caussade, S., De Dios Ortúzar, J., Rizzi, L. I. and Hensher, D. A. 2005. Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B: Methodological,* 39**,** 621-640.

[12] Childers, T. L. and Houston, M. J. 1984. Conditions for a picture-superiority effect on consumer memory. *Journal of Consumer Research,* 11**,** 643-654.

[13] Dhar, R. and Simonson, I. 2003. The effect of forced choice on choice. *Journal of Marketing Research,* 40**,** 146-160.

[14] Eldesouky, A., Mesías, F.J., Elghannam, A., Gaspar, P. and Escribano, M., 2016. Are packaging and presentation format key attributes for cheese consumers?. *International Dairy Journal*, 61**,** 245-249.

[15] Engin, A. and Vetschera, R., 2017. Information representation in decision making: the impact of cognitive style and depletion effects. *Decision Support Systems*, 103**,** 94-103.

[16] Ferrini, S. and Scarpa, R. 2007. Designs with a priori information for non-market valuation with choice experiments: A Monte Carlo study. *Journal of Environmental Economics and Management,* 53**,** 342-363.

[17] Fitzsimons, G. J., Hutchinson, J. W., Williams, P., Alba, J. W., Chartrand, T. L., Huber, J., Kardes, F. R., Menon, G., Raghubir, P. and Russo, J. E. 2002. Non-conscious influences on consumer choice. *Marketing Letters,* 13**,** 269-279.

[18] Green, C. and Tunstall, S. 1999. *A psychological perspective*, Oxford University Press.

[19] Greene, W. H. 2003. *Econometric analysis*, Pearson Education India.

[20] Greene, W. H. 2012. NLOGIT Version 5 Reference Guide. Econometric Software, Inc.: Econometric Software, Inc. Plainview, NY.

[21] Hanley, N., Mourato, S. and Wright, R. E. 2001. Choice modelling approaches: a superior alternative for environmental valuatioin? *Journal of Economic Surveys,* 15**,** 435-462.

[22] Hensher, D., Shore, N. and Train, K. 2005. Households' willingness to pay for water service attributes. *Environmental and Resource Economics,* 32**,** 509-531.

[23] Hensher, D. A., Rose, J. M. and Greene, W. H. 2015. *Applied Choice Analysis*, Cambridge University Press.

[24] Hering, J. G. and Ingold, K. M. 2012. Water resources management: what should be integrated? *Science,* 336**,** 1234-1235.

[25] Hess, S. and Rose, J. M. 2009. Should reference alternatives in pivot design SC surveys be treated differently? *Environmental and Resource Economics,* 42**,** 297-317.

[26] Hetherington, J., Daniel, T. C. and Brown, T. C. 1993. Is motion more important than it sounds?: The medium of presentation in environment perception research. *Journal of Environmental Psychology,* 13**,** 283-291.

[27] Holbrook, M. B. and Moore, W. L. 1981. Feature interactions in consumer judgments of verbal versus pictorial presentations. *Journal of Consumer Research,* 8**,** 103-113.

[28] Howard, J. A. 1977. *Consumer Behaviour in Marketing Strategy.*, New Jersey: Prentice Hall.

[29] Jansen, S., Boumeester, H., Coolen, H., Goetgeluk, R. and Molin, E. 2009. The impact of including images in a conjoint measurement task: evidence from two small-scale studies. *Journal of Housing and the Built Environment,* 24**,** 271-297.

[30] Johnston, R.J., Boyle, K.J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W.M., Hanley, N., Ryan, M., Scarpa, R. and Tourangeau, R., 2017. Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4**,** 319-405.

[31] Jones, M. P. and Hunt, W. F. 2010. Performance of rainwater harvesting systems in the southeastern United States. *Resources, Conservation and Recycling,* 54**,** 623-629.

[32] Kanyoka, P., Farolfi, S. and Morardet, S. 2008. Households' preferences and willingness to pay for multiple use water services in rural areas of South Africa: an analysis based on choice modelling. *Water SA,* 34**,** 715-723.

[33] Keane, M. P. and Wasi, N. 2012. Estimation of discrete choice models with many alternatives using random subsets of the full choice set: With an application to demand for frozen pizza.

[34] Krucien, N., Ryan, M. and Hermens, F., 2017. Visual attention in multi-attributes choices: what can eye-tracking tell us?. *Journal of Economic Behaviour and Organization*, 135**,** 251-267.

[35] Landon, A. C., Kyle, G. T. and Kaiser, R. A. 2016. Predicting compliance with an information-based residential outdoor water conservation program. *Journal of Hydrology,* 536**,** 26-36.

[36] Lanz, B. and Provins, A. 2015. Using discrete choice experiments to regulate the provision of water services: do status quo choices reflect preferences? *Journal of Regulatory Economics,* 47**,** 300-324.

[37] Liu, A., Giurco, D. and Mukheibir, P. 2015. Motivating metrics for household water-use feedback. *Resources, Conservation and Recycling,* 103**,** 29-46.

[38] Louviere, J. J. 2001. Choice experiments: an overview of concepts and issues. *The Choice Modelling Approach to Environmental Valuation***,** 13-36.

[39] Lovett, A., Appleton, K., Warren-Kretzschmar, B. and Von Haaren, C. 2015. Using 3D visualization methods in landscape planning: An evaluation of options and practical issues. *Landscape and Urban Planning,* 142**,** 85-94.

[40] MacInnis, D. J. and Price, L. L. 1987. The role of imagery in information processing: Review and extensions. *Journal of Consumer Research,* 13**,** 473-491.

[41] Makki, A. A., Stewart, R. A., Panuwatwanich, K. and Beal, C. 2013. Revealing the determinants of shower water end use consumption: enabling better targeted urban water conservation strategies. *Journal of Cleaner Production,* 60**,** 129-146.

[42] Marsh, D., Mkwara, L. and Scarpa, R. 2011. Do respondents' perceptions of the status quo matter in non-market valuation with choice experiments? An application to New Zealand freshwater streams. *Sustainability,* 3**,** 1593-1615.

[43] McFadden, D. 1973. Conditional logit analysis of qualitative choice behavior.

[44] Meyerhoff, J. and Liebe, U. 2009. Status quo effect in choice experiments: empirical evidence on attitudes and choice task complexity. *Land Economics,* 85**,** 515-528.

[45] Michalkiewicz, M. and Erdfelder, E., 2016. Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and presentation formats. *Memory & Cognition*, 44**,** 454-468.

[46] Mini, C., Hogue, T. and Pincetl, S. 2015. The effectiveness of water conservation measures on summer residential water use in Los Angeles, California. *Resources, Conservation and Recycling,* 94**,** 136-145.

[47] Molin, E. J. 2011. Conjoint analysis. *The measurement and analysis of housing preference and choice.* Springer, Dordrecht.

[48] Munro, A. and Hanley, N. D. 2001. Information, uncertainty, and contingent valuation. *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries. Oxford University Press, Oxford***,** 258-274.

[49] Norman, R., Viney, R., Aaronson, N.K., Brazier, J.E., Cella, D., Costa, D.S.J., Fayers, P.M., Kemmler, G., Peacock, S., Pickard, A.S. and Rowen, D., 2016. Using a discrete choice experiment to value the QLU-C10D: feasibility and sensitivity to presentation format. *Quality of Life Research*, 25**,** 637-649.

[50] Orzechowski, M., Arentze, T., Borgers, A. and Timmermans, H. 2005. Alternate methods of conjoint analysis for estimating housing preference functions: Effects of presentation style. *Journal of Housing and the Built Environment,* 20**,** 349-362.

[51] Patterson, Z., Darbani, J. M., Rezaei, A., Zacharias, J. and Yazdizadeh, A. 2017. Comparing text-only and virtual reality discrete choice experiments of neighbourhood choice. *Landscape and Urban Planning,* 157**,** 63-74.

[52] Price, G. 2009. Water Conservation Guideline. eThekwini Municipality: eThekwini Municipality

[53] Rizzi, L. I., Limonado, J. P. and Steimetz, S. S. 2012. The impact of traffic images on travel time valuation in stated-preference choice experiments. *Transportmetrica,* 8**,** 427-442.

[54] Ro, T., Singhal, N. S., Breitmeyer, B. G. and Garcia, J. O. 2009. Unconscious processing of color and form in metacontrast masking. *Attention, Perception, and Psychophysics,* 71**,** 95-103.

[55] Rose, J. M. and Bliemer, M. C. 2009. Constructing efficient stated choice experimental designs. *Transport Reviews,* 29**,** 587-617.

[56] Rose, J. M., Bliemer, M. C., Hensher, D. A. and Collins, A. T. 2008. Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B: Methodological,* 42**,** 395-406.

[57] Saldías, C., Speelman, S., Van Huylenbroeck, G. and Vink, N. 2016. Understanding farmers' preferences for wastewater reuse frameworks in agricultural irrigation: lessons from a choice experiment in the Western Cape, South Africa. *Water SA,* 42**,** 26-37.

[58] Samuelson, W. and Zeckhauser, R. 1988. Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty,* 1**,** 7-59.

[59] Scarpa, R., Campbell, D. and Hutchinson, W. G. 2007. Benefit estimates for landscape improvements: sequential Bayesian design and respondents' rationality in a choice experiment. *Land Economics,* 83**,** 617-634.

[60] Scarpa, R., Gilbride, T. J., Campbell, D. and Hensher, D. A. 2009. Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics,* 36**,** 151-174.

[61] Snowball, J., Willis, K. and Jeurissen, C. 2008. Willingness To Pay For Water Service Improvements In Middle-Income Urban Households In South Africa: A Stated Choice Analysis. *South African Journal of Economics,* 76**,** 705-720.

[62] Still, D. and Bhagwan, J. 2008. The Status and Use of Potable Water Conservation and Savings Devices in the Domestic and Commercial Environments in South Africa. *Water Distribution Systems Analysis 2008.*

[63] Syrengelas, K. 2017. Examining text versus visual presentation of choice experiments: Does the presentation method affect consumer preferences and willingness to pay? Master's Thesis, University of Tennessee.

[64] Talati, Z., Norman, R., Pettigrew, S., Neal, B., Kelly, B., Dixon, H., Ball, K., Miller, C. and Shilton, T., 2017. The impact of interpretive and reductive front-of-pack labels on food choice and willingness to pay. I*nternational Journal of Behavioural Nutrition and Physical Activity*, 14**,** 171.

[65] Timmermans, H. and Van Noortwijk, L. 1995. Context dependencies in housing choice behavior. *Environment and Planning A,* 27**,** 181-192.

[66] Train, K. and Wilson, W. W. 2008. Estimation on stated-preference experiments constructed from revealed-preference choices. *Transportation Research Part B: Methodological,* 42**,** 191-203.

[67] Townsend, C. and Kahn, B. E. 2013. The "visual preference heuristic": The influence of visual versus verbal depiction on assortment processing, perceived variety, and choice overload. *Journal of Consumer Research,* 40**,** 993-1015.

[68] van Loo, E.J., Caputo, V., Nayga Jr, R.M., Seo, H.S., Zhang, B. and Verbeke, W., 2015. Sustainability labels on coffee: Consumer preferences, willingness-to-pay and visual attention to attributes. *Ecological Economics*, 118**,** 215-225.

[69] Vásquez, W. F., Franceschi, D. and Van Hecken, G. T. 2012. Household preferences for municipal water services in Nicaragua. *Environment and Development Economics,* 17, 105-126.

[70] Vloerbergh, I., Fife-Schaw, C., Kelay, T., Chenoweth, J., Morrison, G. and Lundéhn, C. 2007. Assessing consumer preferences for drinking water services-Methods for water utilities. *Techneau Project Report D,* 6.

[71] Vriens, M., Loosschilder, G. H., Rosbergen, E. and Wittink, D. R. 1998. Verbal versus realistic pictorial representations in conjoint analysis with design attributes. *Journal of Product Innovation Management,* 15, 455-467.

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

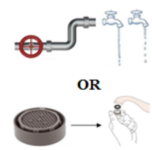
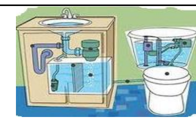
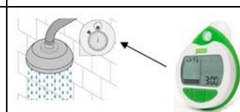
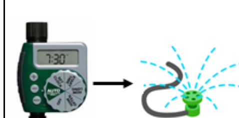
[73] Willis, K. G., Scarpa, R. and Acutt, M. 2005. Assessing water company customer preferences and willingness to pay for service improvements: A stated choice analysis. *Water Resources Research,* 41.

[74] Willis, R. M., Stewart, R. A., Giurco, D. P., Talebpour, M. R. and Mousavinejad, A. 2013. End use water consumption in households: impact of socio-demographic factors and efficient devices. *Journal of Cleaner Production,* 60, 107-115.

[75] Wittink, D. R., Vriens, M. and Burhenne, W. 1994. Commercial use of conjoint analysis in Europe: Results and critical reflections. *International Journal of Research in Marketing,* 11, 41-52.

[76] Wohlwill, J. F. 1976. Environmental aesthetics: The environment as a source of affect. *Human Behavior and Environment: Advances in Theory and Research,* 1, 37-86.

Table 1: Attributes and levels used in the study

Attribute	Description	Attribute Levels	
Kitchen devices 	A typical household uses about 11% of total water use in the kitchen. A standard tap flows at about 8l per minute. Installing water-flow regulators or tap-head aerators saves water and makes a standard tap more efficient by 60%. An efficient dishwasher uses 15l per cycle, using 50% less water than is used in a conventional dishwasher.	**Level 1:** Efficient dishwasher	
		Level 2: Efficient tap	
		Level 3: System collecting used water	
Shower devices 	A typical household uses about 24% of total water in the shower. Shower timers result in shorter showers. Efficient showerheads save 65% of water used in the shower.	**Level 1:** Efficient shower head	
		Level 2: Shower timer	
Toilet devices 	A typical household uses about 25% of total water use in the toilet. Replacing a 12l cistern with a 3l dual cistern saves about 75% of water. An interruptible flush cistern allows users to control how long the toilet flushes. Hippo bags displace water in the cistern and save about 1.2l per flush.	**Level 1:** Dual flush cistern sized 3-6L	
		Level 2: Interruptible flush cistern	
		Level 3: Cistern displacement (hippo bag)	
	A typical household uses about 25% of total water use on garden/outdoor	**Level 1:**	

Garden & outdoor  **devices**	activities. Efficient gardening technologies reduce water use by 30%. These include time-based irrigation control and micro-drip systems. Irrigating gardens using water collected with water tanks also saves water.	Time-based irrigation controller	
		Level 2: Micro-drip systems	
		Level 3: Use harvested rain water	
Monthly water bill 	The average water bill for a household is R450 per month. Installing water-efficient technologies will reduce the monthly water bill by 30%, 50% or 75%.	**Level 1:** R110 **Level 2:** R225 **Level 3:** R315	

Table 2: Example of a text-only choice set

	Status quo	**Option 1**	**Option 2**
Kitchen devices		Efficient dishwasher	System collecting used water
Shower devices		Shower timer	Efficient shower head
Toilet devices		Hippo bag	Dual-flush cistern
Garden/outdoor devices		Time-based irrigation controller	Use harvested rain water
Monthly water bill	R450	R225	R225
YOUR CHOICE			

Table 3: Example of a visual presentation choice set

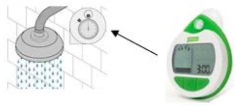
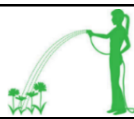
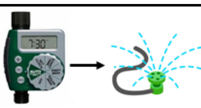
	Status quo	**Option 1**	**Option 2**
			
			
			
			
	R450	R225	R225
YOUR CHOICE			

Table 4: Example of a text-and-visuals choice set

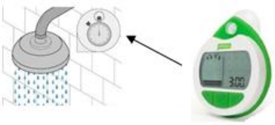
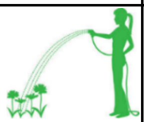
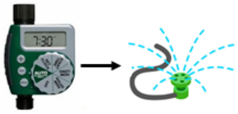
	Status quo	Option 1	Option 2
Kitchen devices 		Efficient dishwasher 	System collecting used water 
Shower devices 		Shower timer 	Efficient shower head 
Toilet devices 		Hippo bag 	Dual-flush cistern 
Garden & outdoor devices 		Time-based irrigation controller 	Use harvested rain water 
Monthly water bill 	R450	R225	R225
YOUR CHOICE			

Table 5: Descriptive statistics of respondents

	Text-only Mean	Visuals-only Mean	Text and Visuals Mean	Pooled data Mean
Number of respondents (N)	***232***	***257***	***405***	***894***
Male respondents (%)	52	64	51	55
Average household size	4	4	4	4
Average age	41	41	45	43
Married respondents (%)	57	45	57	53
Race (%):				
Black	79	89	82	84
White	8	3	13	9
Indian/Asian	9	4	4	5
Coloured	3	4	1	2
Education (%):				
Never attended school	0	0	1	1
Primary	3	1	3	2
High school	68	65	69	67
Certificate	14	21	13	15
Diploma	12	11	10	11
Degree	2	2	4	3
Postgraduate	1	1	1	1
Source of income (%):				
Salaries/wages	57	54	56	56
Business	22	26	20	22
Pension	13	6	17	13
Grants/allowances	3	3	3	3
Other	5	11	3	6

Monthly household income (%):				
<R5 000	34	34	43	38
R5 000 to R10 000	47	50	37	43
R10 000 to R20 000	17	16	19	17
R20 000 to R40 000	3	0	1	1
R40 000 to R60 000	0	0	0	0
>R60 000	0	0	0	0

Table 6: Estimation results on household preferences for water-efficient technology

	Text		**Visuals**		**Text and visuals**	
	Par. Est.	Std. Err	Par. Est.	Std. Err	Par. Est.	Std. Err
Random parameters in utility functions						
KITCHEN	0.008	0.083	0.145**	0.068	0.332***	0.062
SHOWER	0.483***	0.114	-0.065	0.109	-0.204*	0.123
TOILET	-0.075	0.060	0.197***	0.052	-0.052	0.054
GARDEN	0.041	0.074	0.144**	0.073	-0.051	0.065
BILL	-0.012***	0.001	-0.011***	0.001	-0.001***	0.0004
Non-random parameters in utility functions						
ASC	0.0	0.206	0.0	0.427	0.0	0.613
Diagonal values in Cholesky matrix, L.						
NsKITCHEN	0.448***	0.091	0.306***	0.093	0.387***	0.151
NsSHOWER	0.736***	0.159	0.558***	0.151	0.204	0.380
NsTOILET	0.341***	0.087	0.150	0.125	0.001	0.397
NsGARDEN	0.015	0.134	0.346	0.289	0.170	1.234
NsBILL	0.620	0.002	0.211	0.289	0.0004	0.010
Below diagonal values in L matrix. V = L*Lt						
SHOWER:KITCHEN	-0.160	0.183	-0.082	0.184	0.733***	0.151
TOILWT:KITCHEN	-0.122	0.113	0.108	0.103	0.084	0.088
TOILET:SHOWER	0.237*	0.122	0.141	0.122	0.046	0.208
GARDEN:KITCHEN	-0.220**	0.092	-0.238*	0.139	0.026	0.101
GARDEN:SHOWER	-0.065	0.111	-0.723	0.151	0.023	0.264
GARDEN:TOILET	-0.223***	0.082	-0.247	0.319	0.051	4.169

BILL:KITCHEN	0.002	0.002	0.004**	0.002	-0.003***	0.001
BILL:SHOWER	-0.004**	0.002	-0.004*	0.002	0.002	0.003
BILL:TOILET	-0.007***	0.001	-0.004*	0.003	-0.0004	0.033
BILL:GARDEN	-0.0003	0.002	0.003	0.003	-0.001	0.009
Standard deviations of parameter distributions						
sdKITCHEN	0.448***	0.091	0.306***	0.093	0.387***	0.092
sdSHOWER	0.753***	0.155	0.564***	0.152	0.761***	0.217
sdTOILET	0.433***	0.094	0.233**	0.115	0.096	0.155
sdGARDEN	0.321***	0.094	0.493	0.396	0.181	0.111
sdBILL	0.008***	0.001	0.008***	0.003	0.003	0.002
LL Function	-811.4		-897.0		-2459.4	
McFadden Pseudo R^2	0.4		0.4		0.1	
AIC	1664.8		1836.1		4960.9	
BIC	1774.8		1947.8		5082.5	
Number of observations	1392		1515		2324	

Note: ***, ** and * = significance at 1%, 5%, 10% level, respectively. Par Est. = parameter estimates. Std. Err = standard errors

Table 7: Estimates on MWTP for changes in water-efficient devices (in US Dollars)

	Text		**Visuals**		**Text and visuals**	
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
KITCHEN	0.05	0.49	0.90**	0.43	15.59***	5.58
SHOWER	2.84***	0.78	-0.40	0.66	-9.55	6.69
TOILET	-0.44	0.34	1.22***	0.34	-2.45	2.46
GARDEN	0.24	0.44	0.89*	0.48	-2.38	2.94
Wald Statistic	**1.06**		**1.38**		**0.61**	
Prob. from Chi^2	**0.005**		**0.001**		**0.070**	

Note: ***, ** and * = significance at 1%, 5%, 10% level, respectively. Std. Err = standard errors.

Figure 1: Reasons for not having water-efficient technology

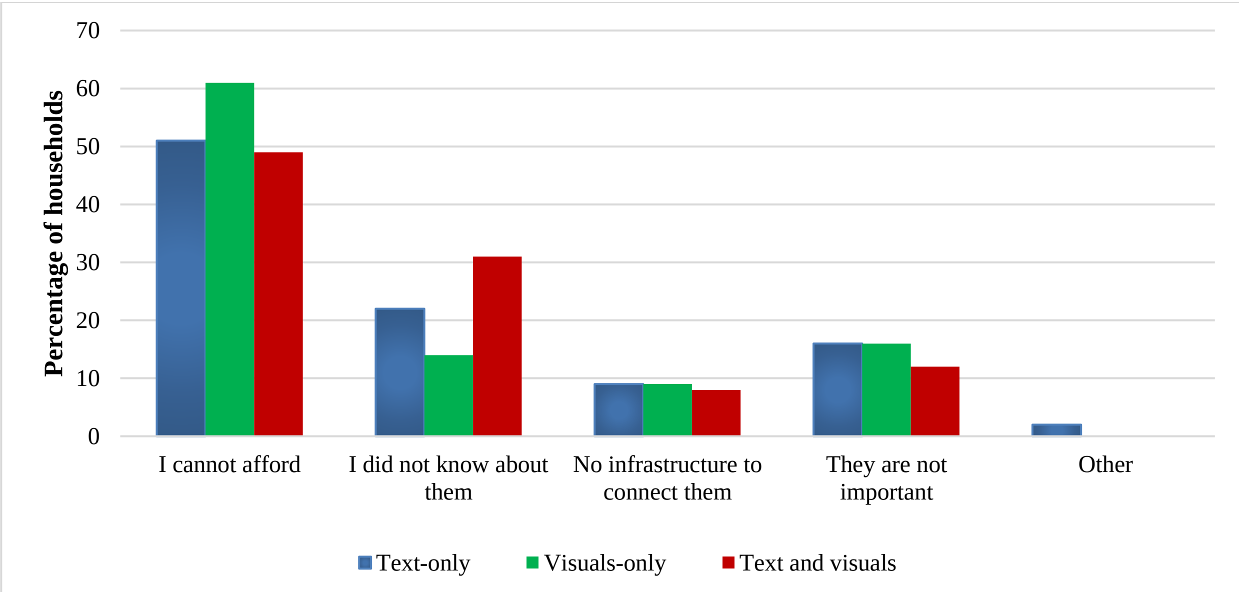


Figure 2: Frequency distribution of stated preference choices made by respondents