Using virtual environments to improve the realism of choice experiments: a case study about coastal erosion management

August 2016

Yvonne Matthews[1], Riccardo Scarpa[1,2,3] and Dan Marsh[1]

[(1)] University of Waikato, N.Z.; [(2)] Durham University, U.K., [(3)] University of Verona, Italy

Forthcoming in *Journal of Environmental Economics and Management*

Abstract

Choice experiment surveys are commonly used to assess the general public's willingness to pay for different levels of environmental quality. However, respondents need to understand what they are valuing or they will make potentially wrong assumptions based on different experiences and frames of reference. Three-dimensional computer generated models or Virtual Environments (VE) have so far seen little use in economics research, probably due to the complexity and cost of developing and delivering them to study participants. The few studies that have used them find that VE are superior to static image presentations in helping people evaluate complex data. For this study we developed virtual environments for a choice experiment about coastal erosion management using free, easy-to-use software and Google Earth© satellite imagery and presented these to respondents as video tours. Our results indicate that the VE treatment reduced choice error, reduced left-right bias and improved respondent engagement and retention when compared with static images. There were also differences in WTP between the two groups.

Keywords

Virtual environment, videos, choice experiment, coastal erosion, scale factor

1. Introduction

In choice experiments, survey respondents are presented with a series of alternative non-market goods and are asked to make trade-offs (state their preference), based on the attributes of the goods which may include environmental quality and cost. For stated preference non-market valuation results to have external validity, participants must be able to comprehend how the changes would affect them in real life and evaluate the alternatives accordingly. Visualisations, such as images, diagrams and maps are usually helpful for conveying complex information to participants (Mathews, Freeman, & Desvousges, 2006). Most visualisations are static, but an alternative is to use dynamic computer-generated 3D environments to convey information about scenarios. Sometimes referred to as virtual environments, the added value of 3-dimensional visualization for the 'evaluability' of unfamiliar scenarios was persuasively argued by Bateman, Day, Jones and Jude (2009). However, virtual environments have rarely been used in non-market valuation - perhaps due to the cost and complexity of developing them. In this paper we describe a relatively cheap and easy method we used to create videos of virtual environments for a choice experiment about coastal development. We use a split-sample and report the differences in models and WTP between the treatment and control group.

2. Literature

It is a well-documented fact that preferences for goods are determined not only by the attributes of a good, but also by the context in which the decision is made. The goal of environmental valuation studies is to determine human response to real environments. Stated preference estimates, however, are based on responses to the information provided by the researcher. Respondent familiarity with an environmental good under valuation is a highly significant predictor of response reliability (Loomis & Ekstrand, 1998) and the information provided is often an inadequate substitute. The choice experiment is an artefactual context that can lack many of the cues that being physically present at the site provides (Fiore, Harrison, Hughes, & Rutström, 2009).

Decision framing effects have been demonstrated in many experimental settings, an early example being the famous study by Tversky and Kahneman (1981). Swait et al. (2002) decomposed the different ways in which context can affect the decision structure: in choice set formation, constraints, evaluation rules and decision rules. Information in a choice task should ideally be presented in a way that minimises "perception confounds" or participant life experiences that influence how they perceive a task (Harrison, Haruvy, & Rutstrom, 2011). Choice experiments often use numerical tables to present information about the alternatives, but these can be either difficult for respondents to comprehend or they may be used in different frames of reference to evaluate them. If the complexity of a task exceeds respondents' median cognitive ability the majority of respondents will make larger errors of judgement than when this is not the case. Consequently, observed choices will appear less consistent (DeShazo & Fermo, 2002). People may also use simplifying strategies to ease choice task execution and not completely process all alternatives and attributes. In addition to having the cognitive abilities to process the task, respondents must also be engaged enough to use their abilities (Bonsall & Lythgoe, 2009). Degrees of engagement, as a broad construct of multiple individual factors, were found by Hess and Stathopoulos (2013) to significantly influence choice consistency.

2.1 Visualisations

Visualisations such as photographs, maps and diagrams were found to both standardise and aid respondent's comprehension in many studies (Mathews et al., 2006). For example, Corso (2001) found that WTP sensitivity to the magnitude of a risk reduction was improved using visual aids rather than just text. Landscapes are particularly difficult to evaluate and even photographs may be inadequate representations for scenic beauty judgements (Daniel & Meitner, 2001). A computer-generated virtual environment is more immersive than static images because visual fields are generated dynamically depending on the simulated viewpoint (Harrison et al., 2011). If the virtual environments are interactive they are referred to as virtual reality (VR). While virtual environments have been used for decades in games and building design, their use in non-market valuation is extremely limited and more recent.

3D visualisations have been developed for several studies (Davies, Laing, & Scott, 2002; Madureira, Nunes, Borges, & Falcão, 2011; Olschewski, Bebi, Teich, Wissen Hayek, & Grêt-Regamey, 2012). However, these were only shown to participants as static images. Jude (2008) found that a combination of 3D visualisation and GIS stimulated more meaningful discussions about coastal planning among planners than did 2D maps, thus improving engagement. Fiore et al. (2009) introduced the use of virtual environments for quantitative analysis of preferences. They found the 3D visualisation generated more accurate beliefs about forest fire risks and recommended it as a way to bring natural field cues into a lab setting. Virtual representation techniques in an area of market research called "information acceleration" have also proven useful for forecasting demand for unfamiliar goods (Urban et al., 1997). Perhaps the first systematic use of virtual environment in a choice experiment was Bateman et al. (2009). Bateman et al. used a fixed flight path so it was not true virtual reality (which is interactive) but they report that the treatment reduced choice error and gain-loss asymmetry in a study of preferences for coastal land-use change.

Creating virtual environments can be a complicated and expensive undertaking. Some researchers such as Davies et al. (2002) used AutoCAD while Olschewski et al. (2012) used Visual Nature Studio. If the virtual environment is to be based on a real location it requires software that integrates with GIS data, such as Arcview© or MapInfo© plugins. Bateman et al. (2009) used a modelling and simulation package called "Terra Vista", while Fiore et al. (2009) used specialised forest fire simulation software. These are all expensive software packages that require specialised skills to use, making their affordability a potential barrier to adoption. However, the effectiveness of visualisation depends more upon their "application in support of communicating the 'concept' and its 'value' to the user" (Hughes, 2004) than the use of state-of-the-art photorealism technology.

The contribution to the literature of this paper is twofold. First, we report a method by which virtual landscapes can be generated using free, easy-to-use software and satellite imagery. We discuss options for delivering the virtual environments to experiment participants via a web survey so that virtual experiments are no longer necessarily restricted to the exclusive settings of expensive computer labs. Second, we contribute to the limited literature about virtual environments in choice experiments by analysing the effect on choice consistency and anomaly reduction.

3. Empirical context

Our study area is beaches of the Coromandel peninsula in the Waikato region of New Zealand. The Coromandel is a steep and hilly peninsula that lies across the Hauraki Gulf from Auckland city. The peninsula is sparsely populated but is a popular holiday destination for residents of the nearby urban areas of Auckland and Hamilton, and to a lesser extent, international tourists. The local population more than doubles during the summer season. There are many coastal landscapes on the peninsula that are considered “Significant Natural Areas” (Graeme, Dahm, & Kendal, 2010) due to their scenic beauty. However, since the 1950s, these coastal areas have been subject to considerable development pressure for holiday accommodation.

Some of the older beachfront developments are now at risk from foreshore erosion and the problem is expected to worsen as sea levels rise. There is conflict between property owners, who want to build seawalls to protect their properties and the council, who have a mandate under the Resource management Act and Coastal Policy Statement to protect natural landscapes and recreation opportunities. Hard coastal defence structures reduce the natural character of a beach, resulting in a loss of sandy foreshore. While there are some short-term options, such as beach nourishment (adding sand), in the long term the main alternative to seawalls is to retreat from the foreshore (remove properties and infrastructure) combined with restoration of the natural dune system.

A qualitative study (Thomson, 2003) found that visitors to the Coromandel peninsula value the natural coastal landscape and recreation opportunities provided. Coromandel tourism expenditure totally $310 million NZD in the year ended March 2014[1]. However, there is a distinct lack of quantitative non-market valuation studies for New Zealand beaches. Some non-New Zealand studies have examined the effects of erosion or sea level rise on beach amenity value including Windle & Rolfe (2014), Whitehead, Poulter, Dumas, & Bin (2009) and Huang, Poor & Zhao (2007). One motivation for the present study is the need to estimate non-market values for the different options for future erosion management on Coromandel beaches, so that both market and non-market costs and benefits can be included in assessments of these options.

4. Method

4.1 Modelling framework

We develop a choice experiment survey to elicit preferences for Coromandel coastal development and estimate marginal utilities. As per random utility theory (RUT McFadden, 1974) we assume the probability of a respondent choosing a scenario is a function of deterministic and random or unobserved components of utility. For reasons of computational tractability we use the logit discrete choice model to develop our analysis. Alternative RUT

[1]http://www.med.govt.nz/sectors-industries/tourism/tourism-research-data/regional-tourism-estimates/regional-summaries

specifications, such as multivariate probit, would also be possible but are outside the scope of this paper.

To reduce the number of choice cards required to achieve statistical significance we require respondents to fully rank alternatives by sequentially selecting their preferred option. These choices are modelled using an exploded logit specification (Lancsar & Louviere, 2008). The utility that person *n* obtains from alternative beach *j* is specified as follows:

$$U_{nj} = \left(ASC_j + \beta_n \mathrm{X_j} + \varepsilon_{\mathrm{nj}}\right) \times \delta_j \tag{1}$$

Where $\mathrm{ASC_j}$ is an alternative-specific constant, $\mathrm{X_j}$ denotes the attribute vector (cost, headland development and erosion protection), β_n is a vector of taste parameters specific to each *n* respondent, ε_{nj} is an i.*i.d* extreme value type one error term, *j* are the alternatives and δ_j denotes whether alternative *j* is available or was already ranked. The β parameters are specified as random with normally distributed density. The unconditional probability of choosing alternative *i* is therefore:

$$P(i) = \int \prod_t \frac{e^{\lambda(\beta' x_i)}}{\sum_j e^{\lambda(\beta' x_j)}} \varphi(\beta|b,W) d\beta \tag{2}$$

Where $\varphi(\beta|b,W)$ is, in our case, a normal density with mean *b* and covariance W. Known as a panel (over *t* choices) mixed (over random β) logit specification, this form allows for taste variation across individuals, unrestricted substitution patterns and correlations in unobserved components across the *t* choices by the same respondent (Train, 2003). The model also includes a scale parameter λ, which cannot be uniquely identified and is inversely related to the variance of the error term in the utility function.

4.2 Measuring the impact on evaluability

Information that is easier to evaluate should reduce "anomalies" in stated preferences or results that conflict with rational choice theory (Bateman et al., 2009). In this study we examine four different indicators of relative evaluability of the alternative scenarios: idiosyncratic choice error, stated choice certainty, frequency of status-quo choices and left-right bias.

4.2.1 Idiosyncratic choice error

RUT includes both random and deterministic components. The random component is a combination of unobserved factors affecting preferences and judgement errors that people make when evaluating the utility of each alternative (Blavatskyy, 2007).

The random or idiosyncratic choice error ε_{nj} can be systematically larger for some individuals and choice situations than others. Choices which are more deterministic have a higher relative scale factor and it is possible to compare the relative scale of pooled datasets such as the VE treatment and control groups (Swait & Louviere, 1993).

The scale factor in stated choice studies is systematically affected by design and respondent-specific factors. In general terms, the greater the gap between choice complexity and respondent's cognitive ability, the higher the idiosyncratic choice error (Caussade, Ortúzar, Rizzi, & Hensher, 2005). In ranking tasks the choice error increases with lower ranks (Ben-Akiva, Morikawa, & Shiroishi, 1992).There is a scarcity of literature about the effects of presentation specifically on choice error. Arentze, Borgers, Timmermans, & DelMistro (2003) found that a presentation with images had no impact on scale factor. On the other hand, Bateman et al. (2009) found that choice variability was lower in the VE treatment group. This paper adds to the limited literature on the effect of presentation formats on scale factor.

Our hypothesis is that the treatment group will have a higher scale factor relative to the control group because the virtual environments make the alternatives easier to evaluate. Using the terms of Swait and Erdem (2007), the videos should increase preference 'discrimination' and reduce the confounding of the preference signal with random error. We parameterise the scale factor and include a treatment dummy variable to test this hypothesis.

4.2.2 Stated Choice Certainty

One way of explicitly accounting for preference uncertainty is to ask people how certain they are about their choice in a follow-up question. An example of this is the Exxon Valdez oil spill study (Carson et al., 1992) in which respondents were asked how strongly they favoured the program. Certainty or follow-up questions in contingent valuation literature appear in several forms and include continuous ratings (Li & Mattsson, 1995), 10-point scales (Champ et al., 1997), 5-point scales (Lundhede, Olsen, Jacobsen, & Thorsen, 2009; Ready, Whitehead, & Blomquist, 1995) or two options "definitely sure" and "probably sure" (Blomquist, Blumenschein, & Johannesson, 2009).The certainty responses may be incorporated directly in the likelihood function (Brouwer, Dekker, Rolfe, & Windle, 2010; Li & Mattsson, 1995) or used to exclude the WTP of uncertain respondents as non-compliant with the assumed fully compensatory behaviour and mitigate hypothetical bias (Champ, Bishop, Brown, & McCollum, 1997)

Similar to idiosyncratic choice error, self-reported certainty has been found to be related to design and individual factors (Lundhede et al., 2009) but individuals with high idiosyncratic error may be more likely to misreport their own certainty (Beck, Rose, & Hensher, 2013). Individuals may also interpret certainty rating scales differently (Loomis & Ekstrand, 1998). So, it is worthwhile examining both the implicit scale factor and stated certainty.

In this study each choice card was followed by a question in which the participant was reminded of their first selection for the card and asked "Do you think this would be your preferred alternative if you really did have to pay?" The response format was a five-point scale comprising "definitely not", "probably not", "maybe", "probably" and "definitely". We test the hypothesis that the VE treatment group will have higher stated certainty on average. We do not include stated certainty as a scale parameter in the logit models because it would be confounded with the VE dummy variable if the hypothesis were correct.

4.2.3 Frequency of status-quo choices

People tend to disproportionately favour an alternative framed as the current situation or status quo (Samuelson & Zeckhauser, 1988). This can be a rational strategy when there are transition costs or the benefits of change are uncertain. Cognitive costs can also cause individuals to favour the status quo because they undertake only partial analysis of the available alternatives. The status quo alternative is advantaged because respondents are familiar with it and understand it better than the alternatives (Scarpa, Willis, & Acutt, 2007). Boxall, Adamowicz, & Moon (2009) found that increased choice complexity leads to increased frequency of the status quo choice, presumably because the analysis costs are higher. Our third hypothesis is that the VE treatment reduces the cognitive cost of alternative evaluation and therefore the magnitude of the status quo advantage. We test this by including a VE times status quo interaction variable in the logit model.

4.2.4 Left-right bias in choice experiments

When alternatives are difficult to evaluate, choices may be influenced more by design factors, such as order of presentation of items, than by the attributes characterizing choice alternatives. Heiner (1983) explained that uncertainty can induce choice behaviour to simple, less sophisticated patterns by the adoption of decision heuristics.

Left-right bias is a systematic result relating to presentation that can arise in choice experiments (Chrzan, 1994). Visually presented items are subject to primacy effects because the first items examined are subject to deeper cognitive processing and establish a standard of comparison (Krosnick, 1999). This implies a left-to-right bias in cultures where individuals read from left to right, and has been reported as an effect (Campbell & Erdem, 2015; Scarpa, Notaro, Louviere, & Raffaelli, 2011). This can be tested, as we do here, by randomising choice profile order and interacting order variables with individual or design-specific parameters such as the VE treatment.

4.3 Equality of willingness-to-pay

Willingness-to-pay (WTP) for a marginal change is the ratio of the attribute coefficient to price coefficient. We use the variance-covariance matrices at convergence and Monte Carlo simulation (Krinsky & Robb, 1986) to approximate the asymptotic sampling distribution of WTP for the video treatment and control groups. Because simulated WTP is not necessarily normally distributed we use the non-parametric Mann-Whitney U test (Mann & Whitney, 1947) to test for equality of mean WTP between both groups.

4.4 Experimental design

The design was kept simple because a virtual environment model and video had to be created for every combination of attributes and levels that affect the visual landscape. The attributes comprise erosion protection, headland development and cost in terms of a tax increase. Headland development is a binary variable, it either occurs or it does not. Thomson (2003) suggests that the number of houses on a headland is irrelevant – if they can be seen from the beach then the natural landscape loses appeal. The erosion options are do nothing, build a seawall or property removal and dune restoration along a specified length of foreshore at risk from erosion. WTP for seawalls may be non-linear with increasing length because of the attitude that any hard protection compromises the natural character of the beach. We use different beach lengths and "at risk" lengths in the design and test for non-linearity in the results for seawalls and dune restoration. Table I shows the attribute descriptions and levels.

Table 1 - Choice experiment attributes and levels

Attribute	Description	Levels
Erosion protection	The beach is *x* km long and *y* km of this has properties at risk from erosion and high waves during storms. The options are to do nothing, remove the front row of properties and restore the nature dune system or build a seawall.	0 = None 1 = Restore dune 2 = Sea wall
Headland	The headland is currently undeveloped and covered with native bush. If development is allowed then houses will be visible in future	0 = No development 1 = Development allowed
Household taxes	Protection of the headland and foreshore require public funding so some of these options will increase your annual rates or taxes by the amount shown	$10 increments from $0 to $100

The status quo option or "current condition" was defined as no erosion protection, development allowed and zero cost. The five other combinations of erosion protection and headland development also appeared on each choice card[2] but in random order to allow for testing of left-right bias. Cost ranged from $0 to $100 per household per year. There were three beaches of different lengths 0.5, 0.8 and 1.2 kilometres and most popular Coromandel beaches fall within this size range. The "the length at risk" to become seawall or restored dune (or neither) was also expressed in kilometres and varied from approximately 30 to 70 per cent of the length of the beach (200 to 800 metres). We tested an orthogonal design in a focus group and obtained prior values with which to generate a Bayesian D-efficient design (Ferrini & Scarpa, 2007) by swapping the cost attribute. This means that cost is no longer orthogonal to the other attributes but this is a more efficient design with which to discern the value of attributes with non-zero priors (Rose, Bliemer, Hensher, & Collins, 2008).

Respondents each received three choice cards, one for each beach of a different length and were asked to rank the six alternatives on each card sequentially. The choice data were modelled as a sequence of five choices from a decreasing set of remaining alternatives, as in Scarpa, Notaro, Louviere and Raffaelli (2011). Respondents were also randomly assigned to a video treatment group or control group. Both groups were presented choice cards with text descriptions of the attribute levels and small images for each alternative. The video group got

[2] Choice cards with fewer alternatives were tested in a focus group but participants disliked not having all the combinations to choose from, even though it made the choice task more complex.

a button to play a video for each alternative while the control group did not. An example of the beach description and choice card can be found in the appendix.

4.4.1 Development of virtual environments

The virtual beach visualisations were developed using Sketchup Make©, a free 3D drawing tool published by Trimble[3]. The first step was to import the terrain and land-cover imagery of a real Coromandel beach from Google Earth©. Realistic yet simple models of houses were added to urban areas by raising building footprints from satellite imagery and draping them with images from Google Streetview©. Sketchup provides a tutorial on how to do this and it only takes a few clicks depending on how many faces the building has.

The study sponsor required that the beaches be unlabelled and did not depict real properties to avoid upsetting property owners about (at this stage) purely hypothetical coastal development. This was not ideal when the goal was to make the experiment as realistic as possible, but we disguised the beaches by draping generic land-cover imagery over easily recognisable landmarks in Google Earth©. The models of buildings were not in their real-world locations and were generic examples of the typical architecture of the region. Participants were informed the beaches were hypothetical, but meant to be representative of beaches in the area. Low-polygon trees and models of people available from the Sketchup 3D Warehouse were dropped into the scene to improve realism of scenarios and provide a sense of scale.

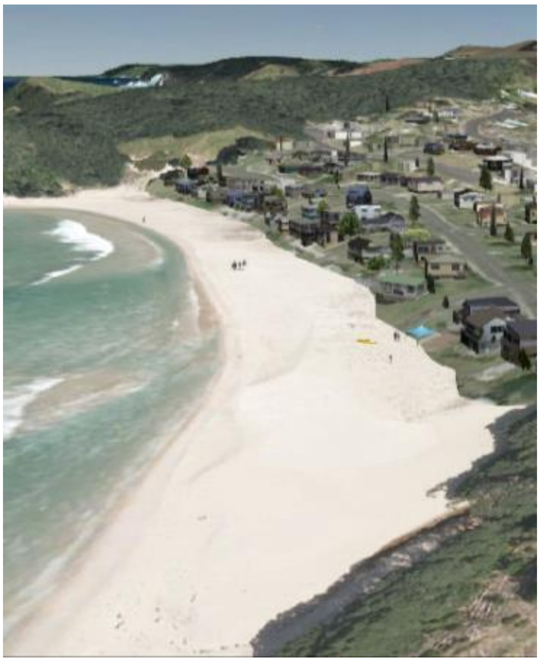

Figure 1 - Bird's eye view of beach with model buildings and props

Seawall models for seawall scenarios were created with a similar height and concrete block texture to that of an existing wall in the Mercury Bay area of the peninsula. For the managed retreat and dune restoration scenarios, the front row of properties was removed and the

[3] http://www.sketchup.com

terrain was raised to form a dune shape. The dune models were draped with a texture from a typical vegetated Coromandel dune. Figure 2 shows still images of the same beach with a seawall or restored dune.

Figure 2 – Beach with status quo, seawall model and restored dune model

For the headland development scenarios, additional buildings were added to the headland at heights such that they appear to be nestled in the herbaceous vegetation. Figure 3 shows a virtual headland with and without houses and a photo of a real headland for comparison.

Figure 3 – Two virtual headlands and a real headland

The virtual beaches were exported as geo-referenced KML files for Google Earth. The Google Earth application or a browser plug-in can be used to view and virtually walk around these models. Interactive virtual environments could be provided to survey participants using the browser plugin but there are three complications: high data usage, compatibility problems with older versions of browsers and difficulty in controlling what participants see. Like Bateman et al. (2009) we traded interactiveness for simplicity and control and recorded pre-defined tours. Each tour lasted 30 seconds, began with a bird's eye view and traversed the length of the beach at the height of a person walking[4]. Tours were embedded in the web survey using the Youtube© API for javascript[5]. The advantage of the Youtube© API is that it provides excellent cross-browser support and the ability to capture events, such as the user starting and stopping the video.

[4] A selection of videos may be viewed at https://youtu.be/1YuvGW4FOSs
[5] https://developers.google.com/youtube/js_api_reference

4.5 The survey instrument and recruitment

The choice experiment was part of a web-based survey developed to gather information about the revealed and stated preferences of domestic visitors to the Coromandel peninsula for features of coastal landscapes. The survey was repeated three times over six months and included questions about their previous and planned beach visits, environmental attitudes, socio-economic variables and the choice experiment questions. In this paper we only report the stated preference choice experiment results from the first wave of the survey. Participants were randomly assigned to the virtual environment video treatment group or no-video control group. Participants in the video group could not reasonably be forced to watch every video for every alternative so they had to click on a video icon to make it start playing. There were three beaches each with a "current state" video and five alternative videos.

Participants were recruited from October to November 2013 from a pre-recruited panel of participants provided by a market research company and a smaller, self-selected sample from online advertisements on Facebook and Google[6]. To take part in the survey respondents had to live in New Zealand and have visited the Coromandel Peninsula in the past year. People who completed the survey were offered either $5 worth of rewards points for panel members, or a $5 Amazon voucher or prize draw for other participants. We advised respondents we would provide aggregate results to a local authority which implied some degree of consequentiality (Vossler, Doyon, & Rondeau, 2012).

Although face-to-face interviews have long been considered the gold standard of stated preference surveys (Arrow & Solow, 1993), this was simply not practical when the sampling frame included the whole of New Zealand. Web surveys exclude households without internet from the sampling frame but this is less of an issue now that 93 per cent of New Zealand households have an internet connection[7]. The use of a pre-recruited panel restricts multiple participations by the same individuals and is an increasingly popular collection mode (Windle & Rolfe, 2011). Other survey modes have different response biases such as that towards older respondents in face-to-face or telephone interviews (Versus Research, 2012).

The survey collection mode may affect responses due to normative or cognitive factors (Dillman, 2011). The physical presence of an interviewer may induce a "social desirability bias" on stated WTP (Groves, Presser, & Dipko, 2004) and provide motivation to put more effort into processing the information. Lindhjem and Navrud (2011) found no evidence of a significant difference in WTP or the degree of satisficing between face-to-face and internet surveys. There may be counterfactual examples, but the fact remains that web surveys are increasingly popular. If virtual environments increase the interest of respondents or reduce the cognitive burden they may be useful for both web surveys and computer-assisted face-to-face interviews.

[6] There were demographic and attitudinal differences between the panel and advertisement samples which are discussed in more detail in a forthcoming technical report.

[7] http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/yearbook/society/technology/connection.aspx

5 Results

5.1 Descriptive statistics

The sample for the choice experiment comprised 1,062 individuals. Table II shows a selection of demographic variables for the samples. The majority of respondents lived in the Waikato or Auckland regions and less than ten per cent were permanent residents of the peninsula. Respondents tended to be older and more highly educated than the general population. Thomson (2003) also found in on-site surveys that visitors to Coromandel beaches were more highly educated than the general population.

Table II - Descriptive statistics

Variable	
Count of respondents	1062
Average age	42
Coromandel resident	0.09
Waikato resident	0.39
Auckland resident	0.36
Female	0.58
Maori ethnicity	0.09
University educated	0.47
Employed full-time	0.48
Preschool children	0.11
Income over $100k	0.33

The average participant made 2.25 trips to visit a Coromandel beach during the six month survey period and spent 6.7 days there in total. The average reported cost of accommodation per night was $11, or $79 excluding people who stayed for free in a private property.

5.2 Virtual environment treatment effects

The majority (81 percent) of respondents in the VE treatment group watched at least one video. The average time spent watching videos was 73 seconds. Figure 4 shows the distribution of the number of videos watched by people in the VE treatment group. There were 18 videos available across the three choice cards but no-one watched more than 6 videos. The number of videos watched declined after the first choice card, perhaps because the scenarios were similar for each beach. Respondents watched an average of 1.4 videos for the first beach, 0.7 for the second and 0.6 for the third.

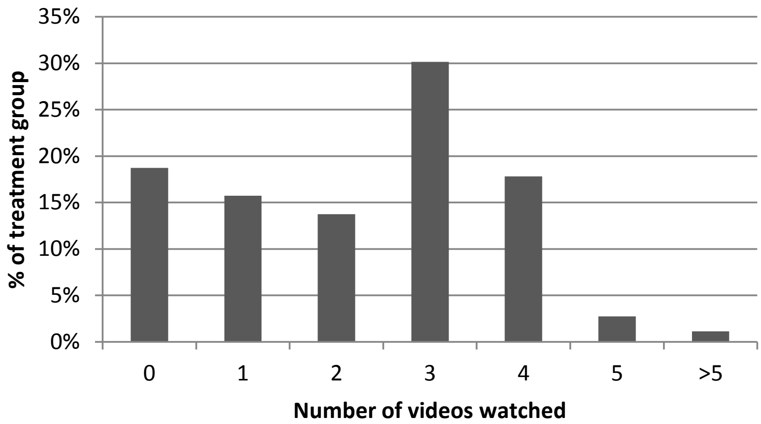

Figure 4 - Number of videos watched by participants in VE treatment group

Respondents in the pilot launch of the survey (n = 136) were asked to give feedback about their survey experience. On a five-point scale of progressively higher enjoyment, the video group gave an average score of 3.8 versus 3.5 for the control group (p = 0.002). The video group were also more likely to agree the survey was "interesting" (82 percent compared with 68 percent for the control group, p = 0.03).

We test whether the video treatment affected stated choice certainty and survey completion rates using Pearson's Chi-squared tests. Table III shows the proportion of respondents in each group who were uncertain of their choices (did not answer "probably" or "definitely" certain), the proportion who completed the post-choice survey questions and retest surveys as well as odds ratios and Chi-square statistics. The video treatment group had a lower rate of stated uncertainty but this was significant only at ten per cent. Differences in survey completion and retest participation, however, are all significant at less than one per cent. The video treatment group were 2.61 times more likely to complete the first survey, 1.83 times more likely to complete the three-month retest and 1.54 times more likely to complete the six-month retest. The virtual environments apparently reduced panel attrition by making the experience more engaging.

Table III – Video treatment effect on stated choice certainty, survey completion and participation

	Uncertain of choice	Completed survey	3mth retest participation	6mth retest participation
Treatment mean	0.22	0.97	0.536	0.403
Control mean	0.26	0.93	0.387	0.304
Odds ratio	0.77	2.61	1.83	1.54
χ_2 statistic	1.86*	9.37***	24.62***	11.63***

* significant at 10%, ** significant at 5%, *** significant at 1%

5.3 Model results

We used maximum simulated likelihood in Biogeme (Bierlaire, 2003) to estimate random parameter logit (RPL) models for the data. The random parameters are normally distributed.

There are dummy variables for headland development, dune restoration and seawall. There are also continuous variables for dune and seawall length. In Table V we present the model results for the video and control groups separately and three pooled models. Pooled model A has a scale parameter for video treatment and interaction variables for video treatment with left-most position, status quo, dune and seawall. Pooled model B only has the scale parameter and pooled model C assumes equal scale.

5.3.1 Log-likelihood ratio tests for pooled models

To test the equality of the pooled and separate models for the video treatment group and control group we use a likelihood ratio (LR) test (Swait & Louviere, 1993). The LR test statistic is calculated as follows:

$$LR = -2\left(LL_{pooled} - (LL_1 + LL_2)\right) \tag{3}$$

Where LL_1 is the final log-likelihood of the model for video treatment sub-sample, LL_2 is that for the control sub-sample, and LL_{pooled} is for the pooled model. The LR statistic has a Chi-square distribution with degrees of freedom equal to the difference in number of parameters. Table IV shows that the LR statistics for pooled models B and C with equal scale exceed the 1 percent critical value so do not explain the data as well as the separate models even after controlling for scale effects. The test statistic for model A is insignificant because the addition of treatment interaction terms sufficiently improves the explanatory power of the pooled model. On the basis of these test results we conclude that results for the two groups are different and the interaction terms discussed in the next section help clarify where they are different.

Table IV - Log-likelihood ratio tests

	Pooled A (λ1 ≠ λ2)	Pooled B (λ1 ≠ λ2)	Pooled C (λ1 = λ2)
LR test statistic	14.51	50.16	88.78
Degrees of freedom	10	16	18
P-value	0.151	<0.001	<0.001

5.3.2 Parameter estimates

The first five parameters account for the effect of position (left-right bias) on participant choice. The choice card is relatively complex with six alternatives and the significant positive coefficients on positions one to five show the right-most item is least likely to be chosen for both groups. The position parameters are relatively smaller for the video group and the VIDEO × POSITION1 parameter in the pooled model is negative, indicating that the video group exhibit less left-right bias. The mean STATUSQUO mean parameter is not significantly different from zero in either group although the random parameter is.

The cost parameter is fixed, rather than random, to avoid the issue of an untenably long upper tail caused by draws that are close to zero (Scarpa, Thiene, & Train, 2008). It is negative and significant for both groups. The HEADLAND parameter – which denotes development is permitted – is also negative and significant for both groups. There is significant taste heterogeneity in the sample as evidenced by the random parameter STDEV_HEADLAND. DUNE, which denotes a restored and planted dune, is significant and positive as is the random parameter STDEV_DUNEDUMMY. DUNELENGTH is not significant but its random parameter is, implying many respondents were insensitive to the size of the restored area.

Preferences for seawalls to protect existing properties are more complicated because some people have positive attitudes towards them and some negative. On average people have a positive WTP but the mean is not significantly different from zero for the video sub-sample. The random parameter STDEV_WALL is significant and slightly larger than the other random parameters indicating wide variation in preferences for seawalls. We tested an alternative latent class specification for seawalls but the overall model fit was poor. Again, the SEAWALLLENGTH mean parameter is insignificant which means people were insensitive to the length of the seawall.

Models A and B have a scale parameter SCALE_VIDEO to test for a systematic difference in scale between the video treatment and control groups. This parameter is positive and significant, which means that the video treatment group exhibits more deterministic choices. This is consistent with the finding that the video sub-sample model has a higher McFadden r-square (0.17 verus 0.11 for the control sample). Model A also has interaction variables for headland, dune, seawall and status quo to test whether the video treatment shifted preferences but none of these are significant.

Table V –Panel mixed logit models

	Pooled A (λ1 ≠ λ2)		Pooled B (λ1 ≠ λ2)		Pooled C (λ1 = λ2)		Video sub-sample		Control sub-sample	
Variable	Coefficient	Z-value	Coefficient	Z-value	Coefficient	Z-value	Coefficient	Z-value	Coefficient	Z-value
POSITION1	0.2980***	9.35	0.5170***	8.63	0.6150***	9.23	0.5060***	5.33	0.7090***	7.60
POSITION2	0.3490***	5.99	0.3510***	6.01	0.4180***	6.30	0.3440***	3.64	0.4950***	5.34
POSITION3	0.2900***	4.97	0.2910***	4.97	0.3440***	5.16	0.2920***	3.06	0.4030***	4.32
POSITION4	0.2450***	4.18	0.2480***	4.21	0.2930***	4.34	0.2710***	2.82	0.3120***	3.28
POSITION5	0.1630**	2.68	0.1640**	2.69	0.1950**	2.78	0.1840*	1.84	0.2170**	2.19
STATUSQUO	0.0199	0.28	-0.0123	-0.26	-0.0077	-0.14	-0.0577	-0.73	0.0584	0.79
COST	-0.0083***	-8.51	-0.0083***	-8.52	-0.0093***	-8.47	-0.0110***	-7.00	-0.0069***	-4.62
HEADLAND	-0.9230***	-12.42	-0.9690***	-17.47	-1.1200***	-18.97	-1.3200***	-13.90	-0.8960***	-11.97
DUNE	0.5100***	4.75	0.5520***	7.62	0.6400***	7.86	0.8210***	6.44	0.5370***	4.89
DUNELENGTH	0.1910	1.01	0.1950	1.59	0.1260	0.90	0.3850*	1.87	0.2030	1.04
SEAWALL	0.2720**	2.34	0.1550**	2.02	0.1710*	1.94	-0.0420	-0.31	0.3090**	2.68
SEAWALLLENGTH	-0.0700	-0.35	0.0791	0.60	0.0663	0.44	0.1090	0.49	-0.0145	-0.08
STDEV_HEADLAND	1.5800***	26.29	1.5900***	26.58	1.8200***	32.82	2.0100***	23.25	1.5500***	21.80
STDEV_DUNEDUMMY	-1.3600***	-22.19	1.3700***	22.45	1.5800***	26.08	1.8100***	18.91	1.3800***	16.32
STDEV_DUNELENGTH	1.1000***	7.10	1.1200***	7.40	1.2500***	6.77	1.2800***	3.87	1.1800***	4.10
STDEV_WALLDUMMY	-1.6500***	-23.11	1.6500***	23.67	1.8800***	27.31	2.3000***	21.32	1.6700***	17.48
STDEV_WALLLENGTH	1.5800***	8.47	1.6300***	9.07	1.8500***	8.84	1.7400***	5.90	1.0600*	1.84
STDEV_STATUSQUO	0.8180***	12.98	0.8160***	12.95	0.9380***	13.95	0.8960***	8.80	0.8530***	9.66
SCALE_VIDEO	1.3100***	5.37	1.3000***	5.44						
VIDEO x POSITION1	-0.1390***	-3.33								
VIDEO x STATUSQUO	-0.0563	-0.63								
VIDEO x HEADLAND	-0.0843	-0.83								
VIDEO x DUNE	0.0747	0.51								
VIDEO x DUNEKM	0.1710	0.70								

VIDEO x WALL	-0.2120	-1.36				
VIDEO x WALLKM	0.2580	0.98				
Observations		16,230	16,230	16230	8135	8095
Log-likelihood		-18,340	-18,358	-18,377	-8875	-9458
Pseudo-R^2		0.141	0.140	0.139	0.170	0.110

* significant at 10%, ** significant at 5%, *** significant at 1%

1

5.3.3 Choice probabilities

Figure 7 reports choice probabilities for the video treatment and control sub-samples. The left bar chart shows how often each alternative is chosen first, when all six alternatives are available. If there were no left-right bias each position would have an equal probability 0.167 of being selected. The video treatment group have a flatter slope and less left-right bias. The second figure shows small differences in the propensity of the video and control groups to different management options.

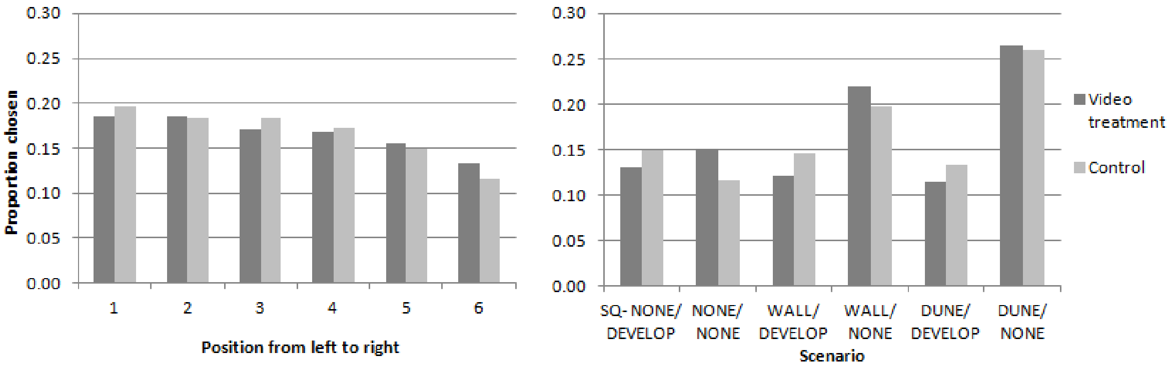

Figure 5 - Choice probabilities for first rank by position and scenario

The model for the video treatment group has slightly better in-sample predictive power and correctly predicts 40 percent of choices across all ranks versus 38 percent for the control group model. By chance alone, we would expect 29 percent correct.

5.3.4 Willingness to pay results

Figure 4 shows box plots of the WTP distributions for headland development and average-length dune and seawall. Visual inspection reveals a large degree of overlap in the distributions for the treatment and the control group with the control group having slightly wider distributions.

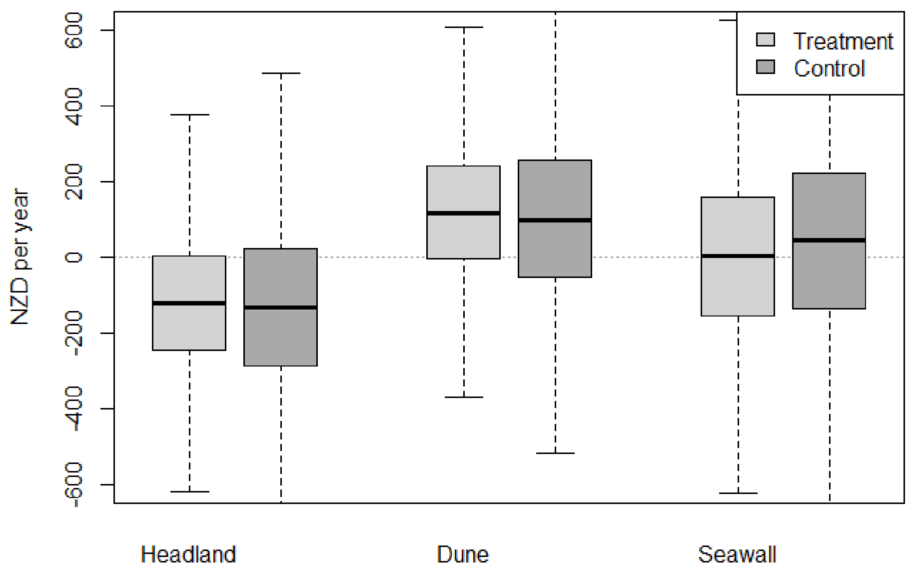

Figure 6 - Boxplot of simulated WTP for video treatment and control groups at mean lengths

In Table VI we present mean WTP, confidence intervals and results of the equality tests for different length dunes and seawalls. Headland development has a mean value of minus $124 for the video group versus minus $138 for the control group. The difference is not significant according to a Mann-Whitney U or Kolmogorov-Smirnov test. The difference in WTP for an 800 metre restored dune versus a 200m restored dune is not large: $43 (47 per cent increase) for the video group and $19 (22 per cent increase) for the control group. Mean WTP for the smallest length of dune restoration is similar for both groups ($91 versus $88) but the video group has a significantly higher WTP for the longest 800m dune. WTP for seawalls is even less sensitive to length and varies by only a few dollars for both groups. The video group has significantly lower mean WTP for seawalls of each length.

Table VI - WTP mean, confidence intervals and tests for differences

		Video treatment		Control		Mann-Whitney	K-S
	Length	Mean	90% C.I.	Mean	90% C.I.	U Z-score	Adj D
Headland	N/A	-124	(-364,114)	-138	(-445,163)	-0.49	1.41
Dune	200m	91	(-126,309)	88	(-186,366)	-1.20	1.62
	400m	106	(-118,332)	96	(-188,387)	0.28	1.68*
	600m	120	(-113,357)	103	(-199,410)	-0.87	2.15**
	800m	134	(-115,387)	107	(-220,441)	-2.25**	2.27**
Seawall	200m	-3	(-278,275)	47	(-281,376)	3.34***	2.95***
	400m	2	(-284,288)	47	(-292,389)	2.54**	2.30**
	600m	2	(-298,304)	46	(-308,401)	2.21**	1.96**
	800m	5	(-316,329)	46	(-330,423)	2.53**	2.21**

6 Discussion

Our finding that WTP is not very sensitive to the length of the dune restoration or seawall is not uncommon in stated preference studies. Also known as embedding or part-whole bias, many researchers have reported evidence of scope insensitivity since it was first demonstrated by Kahneman (1986) and was blamed on the “purchase of moral satisfaction” rather than an economic choice (Kahneman & Knetsch, 1992). Scope insensitivity can be consistent with rational choice theory in situations such as when there are diminishing marginal values for larger area (Rollins & Lyke, 1998), income effects (Randall & Hoehn, 1996) or a lower perceived probability of provision for the larger good (Powe & Bateman, 2004). In the case of seawalls the scope insensitivity may be a result of a perception that any structure on the foreshore reduces the natural character of the whole beach.

In contrast, Carson and Mitchell (1993) argue that scope insensitivity commonly arises when the good or scope are not fully understood by the respondent, referred to as “amenity misspecification” bias. Utility is context-specific (Wilcox, 2011) and choice tasks may define the context imperfectly. If the VE treatment reduces the potential for amenity misspecification it may also increase scope sensitivity. We find a significant difference in WTP for dune restoration for the longest dune in the VE group but this is not strictly speaking a test of scope sensitivity so we are unable to draw any conclusions on the issue.

In section 4.2 we discussed four measures that are affected by respondent difficulties in processing complex information. We use a split sample to investigate whether a VE presentation format affects these measures and present a summary in Table VII. The effect on idiosyncratic error is clear – the scale parameter was higher in both choice experiments in the group with VE treatment. However, it did not appear to make respondents significantly more confident about their choices in terms of stated certainty. The treatment group show less left-right bias, as evidenced by a significant interaction term in pooled model A. The status quo parameter was slightly lower for the treatment group, but the interaction term was insignificant.

We estimated separate and pooled models for the virtual environment treatment and control groups and find that the models are not sufficiently similar even after correcting for scale. The difference is most evident in WTP for seawalls, with the treatment group having a significantly lower mean and a higher proportion with negative values.

Table VII - Summary of video treatment effect

Measure	Test	Result
Survey enjoyment (pilot only)	T-test	Higher
Retest participation	Pearson’s Chi-squared	Higher
Stated choice certainty	Pearson’s Chi-squared	No significant effect
Idiosyncratic error variance	T-test on scale parameter	Lower
Frequency of status quo	T-test on interaction term	No significant effect
Left-right bias	T-test on interaction term	Lower
Willingness-to-pay	Mann-Whitney U, K-S	Lower for seawalls

When there are differences in parameter estimates the question arises as to which values are "right"? Could the videos in fact *alter* preferences rather than elicit them more accurately? It is common knowledge that stated preferences are strongly influenced by framing and presentation effects. However, the literature on stated preference surveys shows that visualisations generally help individuals make more accurate and consistent responses (Mathews et al., 2006). So long as the virtual environment is a fair representation of landscape change then it seems reasonable to assume it will improve the accuracy of elicited values. When respondents can view the landscape from different angles and experience it in a more natural way it reduces the number of potentially wrong assumptions they have to make. The lower choice error variance might also have been due to improved respondent engagement in the VE treatment group. In a climate where people are constantly being asked to do web surveys and response rates are declining, the value of a more engaging survey experience must not be underestimated.

6.2 Areas for future research

This study has shown that a virtual environment can improve the reliability of choice responses in terms of lowering choice error variance. A useful avenue of further research would be to test the effect of VE on external validity, for example by comparing stated preferences for sites with subsequent visits. Virtual environments developed using the tools we describe can be as simple or as complex as the researcher desires. Scenarios of land use change can be represented simply by draping Google Earth terrain with images of a different type of land cover. For a more engaging environment the researcher can add models of buildings, trees, people or other elements from the 3D Warehouse[8]. A VE can also include sound and simple animations created in Sketchup such as a day/night cycle or moving water. The presentation of the virtual environments is not limited to videos of fixed flight paths. Future research could investigate interactive options using the Google Earth browser plugin and API library. This would allow users to freely move around the model while their viewpoints are recorded. Providing an interactive experience does introduce additional technological[9] and methodological complications. More research is required to confirm whether more realistic or interactive virtual environments outperform simple ones, and to what extent and under what circumstances the extra development effort is a worthwhile investment.

7 Conclusion

In this paper we demonstrate a method of developing virtual environments that does not require proprietary GIS data or expensive and complicated modelling and rendering software packages. Nor is the experiment confined to a computer lab setting. The virtual environment can be delivered to web survey participants using free Application Programming Interfaces (APIs) for Google Earth or embedded videos. The treatment has small but statistically significant effects in parameter results and a significant effect on respondent retention. Based on our findings we feel that virtual environments should seriously be considered for use in any non-market valuation study of visible changes to the landscape.

[8] https://3dwarehouse.sketchup.com

[9] Cross-browser compatibility was an issue we encountered when testing the Google Earth plugin and API

8. References

Arentze, T., Borgers, A., Timmermans, H., & DelMistro, R. (2003). Transport stated choice responses: effects of task complexity, presentation format and literacy. *Transportation Research. Part E*, *39*(3), 229–244. http://doi.org/10.1016/S1366-5545(02)00047-9

Arrow, K., & Solow, R. (1993). *Report of the NOAA panel on contingent valuation* (Vol. 58). Retrieved from http://www.cbe.csueastbay.edu/~alima/courses/4306/articles/NOAA%20on%20contingent%20valuation%201993.pdf

Bateman, I. J., Day, B. H., Jones, A. P., & 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.

Beck, M. J., Rose, J. M., & Hensher, D. A. (2013). Consistently inconsistent: The role of certainty, acceptability and scale in choice. *Transportation Research. Part E, Logistics & Transportation Review*, *56*, 81. http://doi.org/10.1016/j.tre.2013.05.001

Ben-Akiva, M., Morikawa, T., & Shiroishi, F. (1992). Analysis of the reliability of preference ranking data. *Journal of Business Research*, *24*, 149–164. http://doi.org/10.1016/0148-2963(92)90058-j

Bierlaire, M. (2003). BIOGEME: a free package for the estimation of discrete choice models.

Blavatskyy, P. (2007). Stochastic expected utility theory. *Journal of Risk and Uncertainty*, *34*(3), 259–286. http://doi.org/10.1007/s11166-007-9009-6

Blomquist, G. C., Blumenschein, K., & Johannesson, M. (2009). Eliciting Willingness to Pay without Bias using Follow-up Certainty Statements: Comparisons between Probably/Definitely and a 10-point Certainty Scale. *Environmental and Resource Economics*, *43*(4), 473–502. http://doi.org/10.1007/s10640-008-9242-8

Bonsall, P., & Lythgoe, B. (2009). Factors affecting the amount of effort expended in responding to questions in behavioural choice experiments. *Journal of Choice Modelling*, *2*(2), 216–236. http://doi.org/10.1016/S1755-5345(13)70011-3

Boxall, P., Adamowicz, W. L. (Vic), & Moon, A. (2009). Complexity in choice experiments: choice of the status quo alternative and implications for welfare measurement. *The Australian Journal of Agricultural and Resource Economics*, *53*(4), 503–519. http://doi.org/10.1111/j.1467-8489.2009.00469.x

Brouwer, R., Dekker, T., Rolfe, J., & Windle, J. (2010). Choice Certainty and Consistency in Repeated Choice Experiments. *Environmental and Resource Economics*, *46*(1), 93–109. http://doi.org/10.1007/s10640-009-9337-x

Campbell, D., & Erdem, S. (2015). Position bias in best-worst scaling surveys: a case study on trust in institutions. *American Journal of Agricultural Economics*, *97*(2), 526–545. http://doi.org/10.1093/ajae/aau112

Carson, R. T., & Mitchell, R. C. (1993). The issue of scope in contingent valuation studies. *American Journal of Agricultural Economics*, *75*(5), 1263–1267.

Carson, R. T., Mitchell, R. C., Hanemann, W. M., Kopp, R. J., Presser, S., & Ruud, P. A. (1992). *A contingent valuation study of lost passive use values resulting from the Exxon Valdez oil spill*. University Library of Munich, Germany.

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

Champ, P. A., Bishop, R. C., Brown, T. C., & McCollum, D. W. (1997). Using Donation Mechanisms to Value Nonuse Benefits from Public Goods. *Journal of Environmental Economics and Management*, *33*(2), 151–162. http://doi.org/10.1006/jeem.1997.0988

Chrzan, K. (1994). Three kinds of order effects in choice-based conjoint analysis. *Marketing Letters*, *5*(2), 165–172. http://doi.org/10.1007/BF00994106

Corso, P. S., Hammitt, J. K., & Graham, J. D. (2001). Valuing mortality-risk reduction: using visual aids to improve the validity of contingent valuation. *Journal of Risk and Uncertainty*, *23*(2), 165–184.

Daniel, T. C., & Meitner, M. M. (2001). Representational validity of landscape visualizations: the effects of graphical realism on perceived scenic beauty of forest vistas. *Journal of Environmental Psychology*, *21*(1), 61–72.

Davies, A.-M., Laing, R., & Scott, S. (2002). Combining visualisation and choice experiments in built environment research. In *Information Visualisation, 2002. Proceedings. Sixth International Conference on* (pp. 785–790). http://doi.org/10.1109/IV.2002.1028870

DeShazo, J. R., & Fermo, G. (2002). Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT*, *44*, 123–143. http://doi.org/10.1006/jeem.2001.1199

Dillman, D. A. (2011). *Mail and Internet Surveys: The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide*. Wiley. Retrieved from http://books.google.co.nz/books?id=d_VpiiWp51gC

Ferrini, S., & Scarpa, R. (2007). Designs with< i> a priori information for nonmarket valuation with choice experiments: A Monte Carlo study. *Journal of Environmental Economics and Management*, *53*(3), 342–363.

Fiore, S. M., Harrison, G. W., Hughes, C. E., & Rutström, E. E. (2009). Virtual experiments and environmental policy. *Frontiers of Environmental and Resource Economics*, *57*(1), 65–86. http://doi.org/10.1016/j.jeem.2008.08.002

Graeme, M., Dahm, J., & Kendal, H. (2010). *Coromandel Peninsula - Ecological Assessment of Natural Character* (Report prepared for Thames-Coromandel Distric Council).

Groves, R. M., Presser, S., & Dipko, S. (2004). The Role of Topic Interest in Survey Participation Decisions. http://doi.org/10.1093/poq/nfh002

Harrison, G. W., Haruvy, E., & Rutstrom, E. E. (2011). Remarks on Virtual World and Virtual Reality Experiments. *Southern Economic Journal*, *78*(1), 87–94.

Heiner, R. A. (1983). The Origin of Predictable Behavior. *The American Economic Review*, *73*(4), 560–595.

Hess, S., & Stathopoulos, A. (2013). Linking response quality to survey engagement: A combined random scale and latent variable approach. *Journal of Choice Modelling*, *7*, 1–12. http://doi.org/10.1016/j.jocm.2013.03.005

Huang, J.-C., Poor, P. J., & Zhao, M. Q. (2007). Economic valuation of beach erosion control. *Marine Resource Economics*, *22*(3), 221.

Hughes, R. (2004). Visualization in transportation: Current practice and future directions. *Transportation Research Record: Journal of the Transportation Research Board*, (1899), 167–174.

Jude, S. (2008). Investigating the Potential Role of Visualization Techniques in Participatory Coastal Management. *Coastal Management*, *36*(4), 331–349. http://doi.org/10.1080/08920750802266346

Kahneman, D. (1986). Comments by Professor Daniel Kahneman. In D. S. B. R.G. Cummings (Ed.), *Valuing Environmental Goods: A State of the art Assessment of the Contingent Valuation Method* (pp. 185–194). New Jersey: Rowman and Allanheld.

Kahneman, D., & Knetsch, J. L. (1992). Valuing public goods: The purchase of moral satisfaction. *JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT*, *22*, 57–70. http://doi.org/10.1016/0095-0696(92)90019-s

Krinsky, I., & Robb, A. L. (1986). On Approximating the Statistical Properties of Elasticities. *The Review of Economics and Statistics*, *68*, 715–719.

Krosnick, J. A. (1999). Survey Research. *Annual Review of Psychology*, *50*(1), 537–567. http://doi.org/10.1146/annurev.psych.50.1.537

Lancsar, E., & Louviere, J. (2008). Estimating individual level discrete choice models and welfare measures using best-worst choice experiments and sequential best-worst MNL. *University of Technology, Centre for the Study of Choice (Censoc)*, 08–004.

Li, C.-Z., & Mattsson, L. (1995). Discrete Choice under Preference Uncertainty: An Improved Structural Model for Contingent Valuation. *Journal of Environmental Economics and Management*, *28*(2), 256–269. http://doi.org/10.1006/jeem.1995.1017

Lindhjem, H., & Navrud, S. (2011). Are Internet surveys an alternative to face-to-face interviews in contingent valuation? *Ecological Economics*, *70*(9), 1628–1637. http://doi.org/10.1016/j.ecolecon.2011.04.002

Loomis, J., & Ekstrand, E. (1998). Alternative approaches for incorporating respondent uncertainty when estimating willingness to pay: the case of the Mexican spotted owl. *Ecological Economics*, *27*(1), 29–41. http://doi.org/10.1016/S0921-8009(97)00126-2

Lundhede, T. H., Olsen, S. B., Jacobsen, J. B., & Thorsen, B. J. (2009). Handling respondent uncertainty in Choice Experiments: Evaluating recoding approaches against explicit modelling of uncertainty. *Journal of Choice Modelling*, *2*(2), 118–147. http://doi.org/10.1016/S1755-5345(13)70007-1

Madureira, L., Nunes, L. C., Borges, J. G., & Falcão, A. O. (2011). Assessing forest management strategies using a contingent valuation approach and advanced visualisation techniques: A Portuguese case study. *Journal of Forest Economics*, *17*(4), 399–414. http://doi.org/10.1016/j.jfe.2011.04.001

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 50–60.

Mathews, K. E., Freeman, M. L., & Desvousges, W. H. (2006). How and how much? The role of information in stated choice questionnaires. In B. J. Kanninen (Ed.), *Valuing environmental amenities using stated choice studies: a common sense approach to theory and practice* (Vol. 8, pp. 111–133). Dordrecht: Kluwer Academic Pub.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*. New York: Academic Press.

Olschewski, R., Bebi, P., Teich, M., Wissen Hayek, U., & Grêt-Regamey, A. (2012). Avalanche protection by forests — A choice experiment in the Swiss Alps. *Integrating Forest Products with Ecosystem Services*, *17*, 19–24. http://doi.org/10.1016/j.forpol.2012.02.016

Powe, N. A., & Bateman, I. J. (2004). Investigating insensitivity to scope: a split-sample test of perceived scheme realism. *Land Economics*, *80*(2), 258–271.

Randall, A., & Hoehn, J. P. (1996). Embedding in market demand systems. *Journal of Environmental Economics and Management*, *30*(3), 369–380.

Ready, R. C., Whitehead, J. C., & Blomquist, G. C. (1995). Contingent Valuation When Respondents Are Ambivalent. *Journal of Environmental Economics and Management*, *29*(2), 181–196. http://doi.org/10.1006/jeem.1995.1040

Rollins, K., & Lyke, A. (1998). The case for diminishing marginal existence values. *Journal of Environmental Economics and Management*, *36*(3), 324–344.

Rose, J. M., Bliemer, M. C. J., Hensher, D. A., & Collins, A. T. (2008). Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B*, *42*, 395–406. http://doi.org/10.1016/j.trb.2007.09.002

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, *1*(1), 7–59. http://doi.org/10.1007/BF00055564

Scarpa, R., Notaro, S., Louviere, J., & Raffaelli, R. (2011). Exploring Scale Effects of Best/Worst Rank Ordered Choice Data to Estimate Benefits of Tourism in Alpine Grazing

Commons. *American Journal of Agricultural Economics*, *93*, 809–824. http://doi.org/10.1093/ajae/aaq174

Scarpa, R., Thiene, M., & Train, K. (2008). Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps. *American Journal of Agricultural Economics*, *90*, 994–1010. http://doi.org/10.1111/j.1467-8276.2008.01155.x

Scarpa, R., Willis, K. G., & Acutt, M. (2007). Valuing externalities from water supply: Status quo, choice complexity and individual random effects in panel kernel logit analysis of choice experiments. *Journal of Environmental Planning and Management*, *50*, 449–466. http://doi.org/10.1080/09640560701401861

Swait, J., Adamowicz, W., Hanemann, M., Diederich, A., Krosnick, J., Layton, D., … Tourangeau, R. (2002). Context Dependence and Aggregation in Disaggregate Choice Analysis. *Marketing Letters*, *13*(3), 195–205. http://doi.org/10.1023/A:1020262503119

Swait, J., & Erdem, T. (2007). Brand effects on choice and choice set formation under uncertainty. *Marketing Science*, *26*(5), 679–697.

Swait, J., & Louviere, J. (1993). The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. \iJournal of Marketing Research, \i30, 305–314.

Thomson, J. (2003). Coastal values and beach use survey report (Technical Report No. TR2003/09). Waikato Regional Council. Retrieved from http://www.waikatoregion.govt.nz/Services/Publications/Technical-reports/2003-technical-reports/

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge Univ Pr.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. http://doi.org/10.1126/science.7455683

Urban, G. L., Hauser, J. R., Qualls, W. J., Weinberg, B. D., Bohlmann, J. D., & Chicos, R. A. (1997). Information acceleration: validation and lessons from the field. *Journal of Marketing Research*, 143–153.

Versus Research. (2012). Community held values of Waikato rivers, lakes and streams in the Waikato and Waipā catchments. *Waikato Regional Council Technical Report*, *2013*(31), 160.

Vossler, C. A., Doyon, M., & Rondeau, D. (2012). Truth in consequentiality: theory and field evidence on discrete choice experiments. *American Economic Journal: Microeconomics*, *4*(4), 145–171.

Whitehead, J. C., Poulter, B., Dumas, C. F., & Bin, O. (2009). Measuring the Economic Effects of Sea Level Rise on Beach Recreation. *Working Papers*.

Wilcox, N. T. (2011). “Stochastically more risk averse:” A contextual theory of stochastic discrete choice under risk. *The Economics and Econometrics of Risk*, *162*(1), 89–104. http://doi.org/10.1016/j.jeconom.2009.10.012

Windle, J., & Rolfe, J. (2014). Estimating the nonmarket economic benefits of beach resource management in southeast Queensland, Australia. *Australasian Journal of Environmental Management*, *21*(1), 65–82. http://doi.org/10.1080/14486563.2013.875953

Windle, & Rolfe, J. (2011). Comparing Responses from Internet and Paper-Based Collection Methods in more Complex Stated Preference Environmental Valuation Surveys. *Economic Analysis and Policy*, *41*(1), 83–97. http://doi.org/10.1016/S0313-5926(11)50006-2

8 Appendix

Figure 7 - Choice experiment 1 beach introduction page

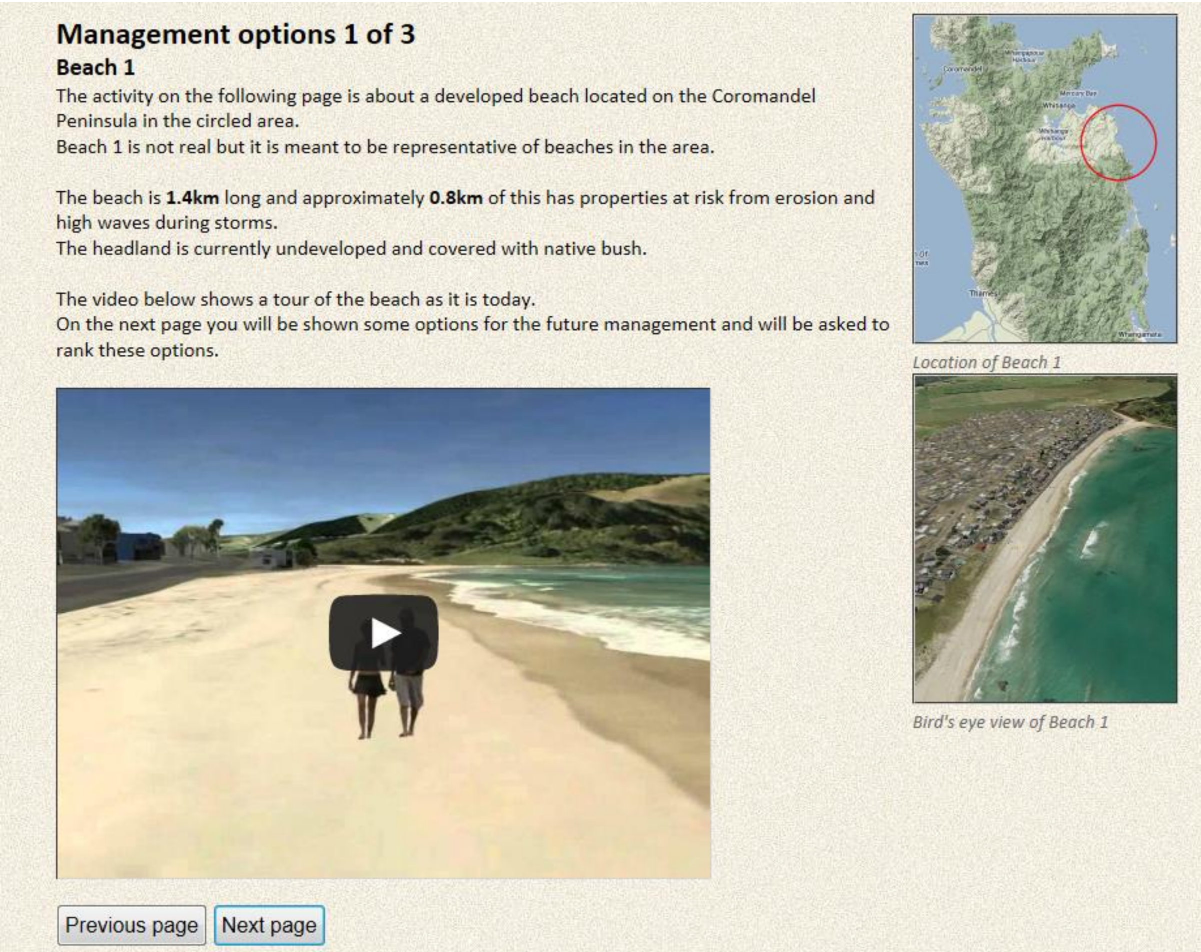

Figure 8 - Choice experiment 1 example card showing mouse-over highlight