Virtual Reality and Scope Sensitivity in a Choice Experiment About Coastal Erosion

Yvonne Phillips and Dr. Dan Marsh

University of Waikato

Contributed paper prepared for presentation at the 59th AARES Annual Conference, Rotorua, New Zealand, February 10-13, 2015

VIRTUAL REALITY AND SCOPE SENSITIVITY IN A CHOICE EXPERIMENT ABOUT COASTAL EROSION

By

Yvonne Phillips and Dr. Dan Marsh, University of Waikato

February 2015

ACKNOWLEDGEMENTS

This research would not have been possible without support and funding from Waikato Regional Council and my supervisors Professor Riccardo Scarpa and Dr. Dan Marsh at Waikato Management School.

ABSTRACT

Choice experiment surveys are a commonly used method to learn the general public's priorities and willingness to pay for different levels of environmental quality. However, results can be sensitive to the framing and presentation of alternative scenarios used in these surveys. Respondents need to understand what they are valuing or they will use simplified heuristics to decide and the results will lack external validity. Visualisation techniques have been found to improve participant's understanding. Virtual reality technologies may be superior to static image presentations but have so far been used in very few studies, probably due to the complexity and cost of developing and delivering the virtual environments to participants. For this study I developed virtual reality scenarios for a web survey about coastal erosion management using free, easy-to-use software and Google Earth satellite imagery. While this method lacks the photorealism that can be obtained from expensive GIS and rendering packages, the results from three different choice tasks show that the virtual reality treatment reduces choice error, improves scope sensitivity and is more interesting for participants when compared with static images.

TABLE OF CONTENTS

1. Introduction 4
2. Literature 4
    2.1 Visualisations 5
    2.2 Measuring the impact on evaluability 5
        2.2.1 Idiosyncratic choice error 6
        2.2.2 Stated Choice Certainty 6
        2.2.3 Left-right bias in choice experiments 7
        2.2.4 Scope insensitivity 8
3. Empirical context 8
4. Method 9
    4.1 The survey instrument 9
    4.2 Development of virtual beaches 10
    4.3 Videos of the whole peninsula 12
    4.4 Experimental design 12
        4.4.1 First Choice task 12
        4.4.2 Second choice task 13
    4.5 Sampling and recruitment 14
5. Results 14
    5.1 Descriptive statistics 14
    5.2 Video treatment 15
    5.3 Choice task 1 model 16
    5.4 Choice task 2 model 18
    5.5 Discussion & conclusion 19
    5.6 Areas for future research 20
    5.7 References 21
    5.8 Appendix 24
        5.8.1 Choice cards 24

1. INTRODUCTION

For stated preference non-market valuation results to have external validity, participants must be able to comprehend and evaluate the alternatives that are presented to them. Visualisations are usually helpful for conveying complex information to participants (Mathews, Freeman, & Desvousges, 2006). Most visualisations such as images, diagrams and maps are static but an alternative is to use dynamic computer-generated 3D environments to convey information about scenarios. Known as virtual reality, computer generated environments are not widely used in stated preference surveys but they may help participant comprehension better than static images.

The purpose of this paper is to answer the question "to what extent can virtual reality improve participant comprehension of a choice experiment task?" The hypothesis is that a virtual reality presentation will systematically affect choice error and sensitivity to the scope and scale of the environmental policy under valuation.

2. LITERATURE

It is a well-documented fact that preferences are determined not only by the attributes of a good but also by the context in which the decision is made. For environmental goods, respondent familiarity with the good under valuation is a highly significant predictor of response reliability (Loomis & Ekstrand, 1998) and the information provided in a stated choice experiment is often an inadequate substitute. The choice experiment is an artefactual context that can lack many of the cues from the field (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. In the present paper I am primarily concerned with whether virtual reality affects the broad validity of choice experiment information.

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 use different frames of reference to evaluate them. If the complexity of a task exceeds respondents' median cognitive abilities the majority 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 such as left-right bias (Chrzan, 1994) or systematically ignore some attributes of the good to ease choice task execution (Scarpa, Gilbride, Campbell, & Hensher, 2009).

2.1 Visualisations

Visualisations such as photographs, maps and diagrams have been found to both standardise and aid respondent's comprehension in many studies (Mathews et al., 2006). Virtual reality (VR) refers to dynamic, computer-generated 3D environments (Harrison et al., 2011) that are more engaging than static images. While computer generated 3D environments have been used for decades in computer games or building design, their use in non-market valuation is extremely limited and more recent.

Davies, Laing, & Scott (2002) and Madureira, Nunes, Borges, & Falcão (2011) created 3D environments of urban and forest landscapes, respectively. However, these were shown to participants as static images. Jude (2008) found that a combination of VR and GIS stimulated more meaningful discussions about coastal planning among planners than did 2D maps. Fiore et al. (2009) introduced the use of VR for quantitative analysis of preferences. They found the VR treatment generated more accurate beliefs about forest fire risks and recommended it as a way to bring natural field cues into a lab setting. Perhaps the first use of virtual reality in a choice experiment was Bateman, Day, Jones and Jude (2009). They report that the VR 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. Davies, Laing, & Scott (2002) used AutoCAD combined with a rendering package to create their visualisations. If the virtual environment is to be based on the real world 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.

The contribution to the literature of this paper is twofold. First, I report a method by which virtual landscapes can be generated using free, easy-to-use software and satellite imagery. I also discuss options for delivering the visualisations 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, I contribute to the limited literature about VR in choice experiments by analysing the effect on evaluability of the choice task by respondents.

2.2 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 I look at five different indicators of relative evaluability of the alternative scenarios: idiosyncratic choice error, stated choice certainty, frequency of status-quo choices, left-right bias and scope (in)sensitivity.

2.2.1 Idiosyncratic choice error

Random utility theory allows for judgement errors or preference imprecision because it includes both random and deterministic components. The probability that individual *n* chooses alternative *i* from set *C* is:

$$P_{i,n} = \frac{\exp(\mu \beta X_{i,n})}{\sum_{j \epsilon C_n} \exp(\mu \beta X_{j,n})}$$

Where *X* are the attributes of the good and μ is the random component that scales all the systematic coefficients and is therefore also known as the scale factor (Swait & Louviere, 1993). The scale factor is inversely related to choice error so more deterministic choices result in a higher scale factor, and vice-versa. The scale factor cannot be identified directly; only compared in relative terms between two or more sets of choice data or sequences of choices by different individual respondents.

Research has shown that the scale factor in stated choice studies is systematically affected by design-specific factors 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). If a design feature, such as the VR treatment, makes the information easier to evaluate then, all else being equal, it should shift the scale factor upwards compared with the control group. In the terms of Swait and Erdem (2007), it would increase preference discrimination.

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 VR presentation treatment group. This paper adds to the limited literature on the effect of presentation formats on scale factor.

2.2.2 Stated Choice Certainty

Stated choice certainty refers to the practice of adding a follow-up question to a stated preference study about the certainty of response. The technique originated in contingent valuation literature where Champ, Bishop, Brown, & McCollum (1997) found it was effective at mitigating hypothetical bias. Certainty scaling questions in the literature appear in several forms and include 10-point scales (Champ et al., 1997), 5-point scales (Lundhede, Olsen, Jacobsen, & Thorsen, 2009) or two options "definitely sure" and "probably sure" (Blomquist, Blumenschein, & Johannesson, 2009).

Similar to idiosyncratic choice error, self-reported certainty has been found to be related to design and individual factors. Stated choice certainty has been found to be associated with lower choice error, as might be expected (Lundhede et al., 2009), but individuals with high idiosyncratic error may be more likely to misreport their own certainty (Beck, Rose, & Hensher, 2013) so it is worthwhile examining both indicators.

2.2.3 Frequency of status-quo choices

People tend to disproportionately favour an alternative framed as the current situation or status quo (Samuelson & Zeckhauser, 1988). One reason for this status quo bias arises when individuals undertake only partial analysis of the available alternatives. The status quo alternative when 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 choice of the status quo. If the VR videos make alternatives easier to evaluate this may reduce status quo bias

2.2.4 Left-right bias in choice experiments

The literature on scale parameter and choice error assumes that a larger complexity-cognitive ability gap makes choices more random giving rise to less discriminating preferences. An alternative theory is that choices may actually be *more* deterministic in the face of uncertainty, but influenced more by design factors than by the attributes of the choice alternatives. Heiner (1983) explained that uncertainty can constrain choice behaviour to simple, less sophisticated patterns by the adoption of heuristics.

If there are a lot of attributes to process the respondent may systematically ignore some to simplify the decision (Scarpa et al., 2009), or some attributes might be merged together when they can be measured by the same metric (Greene & Hensher, 2010). In this study there are only three attributes, so the potential for attribute non-attendance is limited and none share the same metric. Left-right bias is another systematic result 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, which has been reported as an effect (Campbell & Erdem, 2014; Scarpa, Notaro, Louviere, & Raffaelli, 2011). This can be tested by randomising choice profile order and interacting order variables with individual or design-specific parameters such as the VR treatment.

2.2.5 Scope insensitivity

Scope insensitivity refers to the phenomenon in stated preference surveys where respondents appear to have the same WTP for goods that differ largely in scope or inclusiveness. Also known as embedding or part-whole bias, many researchers have reported evidence of scope insensitivity since it was first demonstrated by Kahneman (1986). Scope insensitivity conflicts with rational choice theory because WTP should be higher for a larger or more comprehensive good rather than a subset of the good – assuming demand is non-satiated.

There are many reasons given in the literature for scope insensitivity. One reason is that the good may be poorly described or the scope meaningless to the respondent (Mitchell & Carson, 1989). Therefore, greater scope sensitivity can be considered evidence of improved evaluability if other factors are controlled for.

Rolfe and Wang (2011) categorise the issue into geographic, policy and attribute scope insensitivity. In this paper I report the effect of a VR treatment on data from two choice experiments, one of which has different levels of geographic scope and the other with varying levels of attribute scope.

By investigating whether VR is useful for improving scope sensitivity I extend the work of Bateman et al. (2009), which focussed on gain/loss asymmetry.

3. EMPIRICAL CONTEXT

The subject of this empirical research is 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 rare scenic beauty. However, since the 1950s, these coastal areas have been subject to considerable development pressure.

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, in the long term the main alternative to seawalls is managed (or unmanaged) retreat 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. One of the goals of this study is to provide quantitative non-market values for the alternative erosion management options.

Coastal scientists at Waikato Regional Council developed coastal hazard lines as a tool to manage development and plan for erosion risk[1]. The Primary Development Setback (PDS) delineates developed coastline at risk from natural beach erosion under current conditions. The Secondary Development Setback (SDS) delineates additional land at risk under expectations of sea level rise. These hazard lines are used to define “high” and “medium” risk in the second choice task.

4. METHOD

4.1 The survey instrument

A survey was developed to gather information about revealed and stated preferences of domestic visitors to the Coromandel peninsula. The survey included questions about their previous and planned beach visits, environmental attitudes, socio-economic variables, and the choice experiment questions. The survey was conducted over three waves spanning a period of six months so as to gather additional information about recent beach trips, preference stability and stated intentions versus actual visit behaviour.

The first wave of the survey included a choice experiment about four Coromandel beaches of different sizes with different development scenarios comprising headland development, a seawall or managed retreat and dune restoration. The third wave of the survey included a choice task about erosion protection options for two levels of geographic scope – Mercury Bay area and the whole peninsula.

Each participant was randomly assigned to a video or no-video treatment group for the choice experiment. The video treatment group had the option of playing videos of a virtual reality representation of each scenario.

[1] http://www.waikatoregion.govt.nz/Services/Regional-services/Regional-hazards-and-emergency-management/Coastal-hazards/Development-setback-recommendations-Coromandel-beaches/

4.2 Development of virtual beaches

The virtual beaches scenarios were developed using Sketchup Make, a free 3D drawing tool published by Trimble[2]. The first step was to import the terrain and land-cover imagery of a real Coromandel beach from Google Earth. At the time of this study there was no Google Earth 3D imagery available for New Zealand so models of buildings had to be created from scratch. Simple, low polygon models were created by raising buildings' footprints from satellite imagery and draping these objects with images from Google Streetview.

The study sponsor required that the beaches be unlabelled and did not depict real properties to reduce the risk of upsetting property owners. This complicated the experiment, but the four beaches were disguised by draping generic land-cover imagery over easily recognisable landmarks. The models of buildings were not in their real-world locations and are 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. To improve realism of scenarios and provide a sense of proportions, low-polygon trees and models of people (avatars) from the Sketchup 3D Warehouse were added.

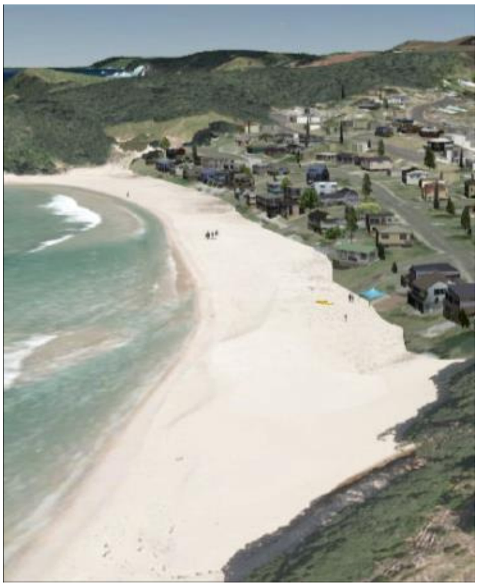

Figure 1 - Bird's eye view of beach with model buildings and props

To create scenarios with seawalls I created models of walls with similar height and concrete block texture to that of an existing seawall in the area. For the managed retreat and dune restoration scenarios I removed the front row of properties, raised the terrain to form a dune shape and draped it with a texture from a natural, vegetated dune. Figure 2 shows still images of the same beach with a seawall or restored dune.

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

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

For the headland development scenarios I added additional buildings to the headland at heights such that they appear to be nestled in the herbaceous vegetation. Figure 3 shows a headland with and without houses.

Figure 3 – Beach headland with and without model buildings

I exported the completed virtual scenarios as geo-referenced KML files for Google Earth. The Google Earth application or a browser plugin can be used to view and virtually walk around these models. Interactive virtual reality models could be provided to survey participants using the browser plugin but there are three complications: high data usage, compatibility problems with older browsers and difficulty in controlling what participants see. Similar to Bateman et al. (2009), I sacrificed interactiveness for simplicity and control by recording 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.

I recorded videos of the tours and embedded these in the web survey using the Youtube API for javascript[3]. The advantage of the Youtube API is that it provides excellent cross-browser support and the ability to record events, such as the user starting and stopping the video.

[3] https://developers.google.com/youtube/js_api_reference

4.3 Videos of the whole peninsula

The second choice task did not require the construction of virtual models. The purpose of the videos was simply to communicate the location and scope of coastline at risk from erosion. The level of risk was calculated by coastal scientists and provided in the form of a shape-file with colour-coded lines running parallel to the shore (see Figure 4). I imported this shape-file to Google Earth and recorded fly-by tours of the area.

Figure 4 - Image from Google Earth tour showing colour-coded levels of risk to property

4.4 Experimental design

4.4.1 First Choice task

The first choice task was kept relatively simple because a video had to be created for every combination of attributes in the design. The design included only two attributes (headland development and foreshore management) plus a cost in terms of a tax increase. Headland development is a binary variable, it either occurs or it does not. A qualitative study of the Coromandel (Thomson, 2003) suggests that the number of houses on a headland is irrelevant – if they can be seen from the beach then it loses natural landscape appeal. The foreshore management options are either a seawall or managed retreat and dune restoration along the developed length of the beach.

Each of the six combinations of attributes appeared on each choice card[4]. Cost ranged from $0 to $100 per year and the ten-block design was optimised for D-efficiency with Bayesian priors obtained

[4] Choice cards with fewer alternatives were tested in a focus group but participants disliked not having all the options to chose from even though it made the choice more complex.

from a focus group (Ferrini & Scarpa, 2007; Sandor & Wedel, 2001). Each participant received one choice card for each of the four beaches and was asked to rank the 6 alternatives sequentially.

4.4.2 Second choice task

The second choice task had cost plus three attributes representing the total proportion of developed coast "at risk" that will be protected by a seawall. The lengths at risk are classified according to level of risk and whether it is private or public property at risk. Table 1 shows the total kilometres of coastline in each category. "High risk – private" is where private properties are in front of (or overlapping) the Primary Development Setback (PDS) line. "Medium risk – private" is where private properties are in front of the Secondary Development Setback (SDS) but not the PDS. "Public risk" refers to the length of coastline where public roads are in front of either the PDS or SDS.

Table 1- Kilometres of developed coastline protected and at risk

Geographic scope	High risk - private	Medium risk - private	Public risk	Existing seawalls	Total developed length
Mercury Bay area	7.9	1.6	6.1	1.5	23.5
Entire peninsula	14	3.6	8.3	2.1	39

There were two geographic areas used in the choice experiment, Mercury Bay area (Figure 5) and the entire peninsula as per the Thames-Coromandel district boundary. The Mercury Bay area has several popular beaches and 60 per cent of the developed coastline and 57 per cent of the existing length of seawalls. Approximately two-thirds of visitor nights are spent in the area. Average willingness to pay is expected to be higher for the whole peninsula compared with that for the smaller area if sensitivity to scope is to be corroborated.

Figure 5 - Map of Coromandel Peninsula and Mercury Bay area (shaded)

The "cost" attribute for this choice task was a change in average nightly accommodation cost for the area. The rationale is that costs incurred protecting or repairing property or infrastructure would be passed on from accommodation providers to visitors. Similar to the first choice task, the design was optimised by minimizing the expected D-error using Bayesian priors for the utility coefficients.

4.5 Sampling and recruitment

Participants were recruited using online advertisements on Facebook and Google (the self-selected sample) and from a pre-recruited panel of participants provided by a market research company. 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. Data was collected in three waves in November 2013, February and June 2014.

5. RESULTS

5.1 Descriptive statistics

The sample for wave one of the survey comprised 822 individuals. There was considerable attrition over the six month period and only 380 completed wave three. Table 2 shows a selection of demographic variables for the samples. The majority of participants live in the Waikato or Auckland regions and less than ten per cent are permanent residents of the peninsula. Participants tended to be older and more highly educated than the general population but Thomson (2003) found that visitors to Coromandel beaches were more highly educated than average.

Table 2 - Descriptive statistics

Variable	Wave 1	Wave 3
Count	822	380
Age	42	44
Coromandel resident	0.09	0.02
Waikato resident	0.39	0.36
Auckland resident	0.36	0.41
Female	0.58	0.63
Maori ethnicity	0.09	0.04
Post-school education	0.71	0.76
Employed full-time	0.48	0.50
Preschool children	0.11	0.13
Income over $100k	0.33	0.27

The average participant made 2.25 trips to visit a Coromandel beach during the 6 month survey period and spent 6.7 days there in total. The average cost of accommodation per night was $11, or $79 excluding people who stayed for free in a private property.

5.2 Video treatment

Participants were randomly assigned to the videos or no videos treatment group. However, participants in the video treatment group could not be forced to watch every video for every alternative. They had to click on a video icon to make it start playing. Table 3 shows the average proportion of videos started, completed, and total time played by participants. For the first choice task there were three beaches each with a "current state" video and five alternative videos. The number of videos watched declined after the first choice card, perhaps because the scenarios were the same for each beach. On average, participants clicked on 60 per cent of the status quo videos and 10 per cent of the alternative scenario videos.

For the choice task in wave 3 there was only one video per choice card, depicting the areas of coast at risk of erosion. On average around 60 per cent of people clicked on the video for each area.

Table 3 - Videos watched by participants in treatment group

	% of videos started	% of videos finished	% of time played
Choice task wave 1			
Status quo videos	60%	46%	49%
Alternative videos	10%	4%	6%
Choice task wave 3			
Mercury Bay	59%	52%	56%
Whole peninsula	63%	52%	53%

Table 4 shows the choice task completion rates and the proportion of people who said they were "unsure" about their choice. The video treatment group were slightly more likely to complete the survey and more certain of their choices. However, these differences are not statistically significant.

Table 4 - Video treatment, survey completion and stated choice certainty

Video treatment	% Completed survey	% Unsure of choice
No	91.8%	25.5%
Yes	92.5%	22.0%

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 VR treatment group gave an average score of 3.8 while the control group were 3.5. The VR treatment group were also more likely to agree the survey was “interesting” (79 per cent compared with 63 per cent for the control group).

5.3 Choice task 1 model

All models were estimated using Biogeme (Bierlaire, 2003). The model presented in Table 5 is a heteroskedastic panel mixed logit, a flexible specification that allows for both heteroskedasticity and preference heterogeneity in the choice data (Fiebig, Keane, Louviere, & Wasi, 2010). The McFadden pseudo r-square is high at 0.83.

The first nine parameters account for the effect of position (left-right bias) on participant choice. The choice card is relatively complex with six alternatives, and the high value for POSITION1 shows that the left-most item was significantly more likely to be chosen. The interactions with FACEBOOK and GOOGLE indicate that those recruited using the online advertisements were less likely to show a left-to-right bias. The video treatment does not appear to have a significant effect on left-to-right bias.

The DUNEDUMMY variable indicates that the alternative has a restored and planted dune, while DUNEKM is the length of the restored area. On average, participants were insensitive to the length of the restored dune and the difference in WTP for a longer dune is insignificant. But the DUNEKM × VIDEO interaction is positive and significant, indicating that the video treatment group were more sensitive to the length. The SEAWALLDUMMY variable is negative, indicating that most participants view this as a negative feature of a beach. Participants were not sensitive to the length of the seawall; not even in the video treatment group.

The VIDEO × COST interaction is positive but not significant so the VR treatment did not significantly increase average willingness to pay. HIGHINCOME × COST is to account for an income effect but is insignificant. The WAIKATO × COST interaction is positive and highly significant, which means that Waikato region residents have higher WTP than people who live further away. I also tested a continuous variable for travel distance, but it was not as significant as the dummy variable.

The mean STATUSQUO parameter is close to zero but the significant random parameter shows there is heterogeneity around the mean. The STATUSQUO x VIDEO parameter is insignificant. The video treatment did not seem to affect respondent propensity to choose the status quo.

The scale shifting parameters include post-school education and rank order, which are known to affect choice error (Ben-Akiva, Morikawa, & Shiroishi, 1992). Post-school education is associated with a higher scale factor and increasing rank order with lower scale, as expected. There was no detectable scale effect for choice card order, and is not necessarily expected over such a small number of cards (Arentze et al., 2003).

The scale factor is parameterised to test for the effect of the VR treatment and to control for design and individual-specific factors that are expected to affect scale. The SCALE_FACEBOOK and SCALE_GOOGLE parameters control for the different demographic profiles of Facebook and Google participants compared with pre-recruited panel participants. SCALE_YOUNGKIDS is a dummy variable for having pre-school children in the household and is intended as a proxy for being time-poor or having frequent distractions. The effect on scale is negative and much more significant than other variables tested such as age, employment status or the presence of school-age children. The SCALE_VIDEO parameter is positive and significant at the ten per cent level, indicating that the video treatment reduced choice error.

Table 5 – Heteroskedastic panel mixed logit model for individual beach choice data

Variable	Coefficient	Z-value
POSTION1	1.8200***	14.21
POSTION2	0.1650***	3.40
POSTION3	0.1180**	2.51
POSTION4	0.1080**	2.27
POSTION5	0.0992**	2.05
POSITION1 x EDUCATION	-0.4290***	-4.14
POSITION1 x FACEBOOK	-0.4270**	-2.17
POSITION1 x GOOGLE	-0.9130***	-6.99
POSITION1 x VIDEO	0.0418	1.20
COST	-0.0008	-1.05
DUNE	-0.0354	-1.09
DUNEKM	0.0065	0.09
DUNEKM x VIDEO	0.1520**	2.29
SEAWALLDUMMY	-0.0823**	-2.46
SEAWALLKM	-0.0366	-0.50
SEAWALLKM x VIDEO	0.0472	0.74
HEADLANDDUMMY	-0.0939***	-5.40
VIDEO x COST	0.0008	1.00
HIGHINCOME x COST	0.0003	0.45
WAIKATO x COST	0.0034***	3.59
STATUSQUO	-0.0553	-1.41
STATUSQUO x VIDEO	0.0392	0.84
STDEV_SEAWALL	-0.2260***	-6.85
SCALE_EDUCATION	0.5300***	4.29
SCALE_FACEBOOK	-0.4770**	-2.26
SCALE_GOOGLE	0.2470*	1.98
SCALE_RANK2	-0.0295	-1.64
SCALE_RANK3	-0.1220***	-3.23
SCALE_RANK4	-0.1550***	-3.75
SCALE_RANK5	-0.2170***	-4.42
SCALE_YOUNGKIDS	-0.1680***	-5.33

SCALE_VIDEO	0.0890*	1.89
Number of observations		15580
Number of individuals		779
Log-likelihood		-3411.53
Pseudo-R^2		0.834
Bayesian information criteria		7131.98

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

5.4 Choice task 2 model

The estimates for the heteroskedastic panel mixed logit model for the second choice task are reported in Table 6. Unlike in the first choice task there was no significant left-to-right bias, perhaps because there were only three alternatives rather than six.

HIGHRISK is a dummy variable indicating that all high risk lengths of private property in the specified area are protected with a seawall. MEDRISK means that all medium risks lengths get a seawall. PUBLIC means that all public infrastructure (mostly roads) get seawall protection. The coefficients are significantly larger for high risk areas.

The survey included a question about general attitudes towards the construction of seawalls for protection. Almost half of participants were against the building of seawalls in general, while 16 per cent thought they were a good thing and the remainder were unsure. NEG and POS are dummy variables for attitude to seawalls. The interactions show that people with a negative attitude have a negative WTP for seawalls, as expected. The exception is for public property, where the WTA for people with a negative attitude is not significantly different from zero.

The payment vehicle in this task is an average increase in accommodation cost across the peninsula. Interaction variables are included for (i) the actual amount participants spent on accommodation while in the Coromandel during the past year, (ii) high income (over $100,000 per year) and (iii) the video treatment dummy; none of which have coefficient estimates found to be statistically significant.

The WHOLE dummy variable indicates the choice card was for the whole peninsula, rather than the sub-region Mercury bay. Each participant was informed, during the introduction, that they would receive a choice card for both areas. The order was random and did not have a significant effect on WTP (WHOLEFIRST × PRICE). For the average participant there was no difference in WTP between the two geographic scopes. The VIDEO × WHOLE × COST interaction coefficient is positive and significant at the ten per cent level, indicating that the video treatment group had a higher WTP for the larger area and were therefore more sensitive to scope.

The status quo bias is positive and significant. Similar to the first choice task, the STATUSQUO x VIDEO interaction is insignificant.

Scale-shifter parameters are included for education, preschool children in the home and the video treatment, similar to the model for the first choice task. The video treatment had a large and positive effect on the scale parameter. The coefficient estimates for SCALE_EDUCATION and SCALE_YOUNGKIDS have the expected sign, but are insignificant.

Table 6 – Heteroskedastic panel mixed logit model for choice task 2

Variable	Coefficient	|Z-value|
LEFT	-0.0305	-1.44
PRICE	-0.0037***	-3.16
HIGHRISK	0.1490***	3.06
HIGHRISK × NEG	-0.3930***	-3.22
HIGHRISK × POS	0.1140**	2.06
MEDRISK	0.0312	0.97
MEDRISK × NEG	-0.1200*	-1.86
MEDRISK × POS	0.1210***	2.48
INFRASTRUCTURE	0.0898**	2.10
INFRASTRUCTURE × NEG	-0.0100	-0.15
INFRASTRUCTURE × POS	0.1120**	1.96
ACCOMSUM × PRICE	-0.000002	-1.36
HIGHINCOME × PRICE	0.000339	0.41
VIDEO × PRICE	0.0007	1.42
WHOLE × PRICE	-0.0011	-1.48
WHOLEFIRST × PRICE	-0.0011	-1.41
VIDEO × WHOLE × PRICE	0.0017*	1.77
SCALE_EDUCATION	0.3600	1.27
SCALE_VIDEO	0.6440***	2.96
SCALE_YOUNGKIDS	-0.1380	-0.46
STATUSQUO	0.1130**	2.10
STATUSQUO x VIDEO	-0.0334	-0.60
STDEV_STATUSQUO	0.2480**	2.27
Number of observations		760
Number of individuals		380
Log-likelihood		-724.13
Pseudo-R^2		0.105
Bayesian information criteria		1558.56

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

5.5 Discussion & conclusion

The literature on stated preference surveys shows that visualisations help individuals make more accurate and consistent responses. Scenarios that involve changes to a landscape are particularly difficult to evaluate without visualisations. Virtual reality is more engaging than static images are and it allows respondents to view the changes from different angles and experience it in a more natural way. Despite the benefits identified by the few studies that have used VR, the technique has not gained widespread use in non-market valuation. This is probably due to a combination of cost and lack of necessary skills amongst the community of researchers in this field.

In this study I demonstrate a method of developing virtual landscapes 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 Youtube videos.

I use a split sample to investigate whether a VR presentation format improves the reliability of choice experiment responses in a sample of visitors to the Coromandel Peninsula. The results show that the VR treatment reduces idiosyncratic choice error, thereby improving preference discrimination. It also improves sensitivity to attribute scale for one of the two non-binary attributes employed to describe policy outcomes. Finally, the VR treatment also improves sensitivity to geographic scope between the whole peninsula and a sub-region of the peninsula.

Furthermore, from the respondent's perspective, the VR treatment made the survey experience more enjoyable and interesting. In a climate where people are constantly being asked to do web surveys and response rates are declining, the value of a more enjoyable experience must not be underestimated. VR should be seriously considered as a presentation format for any non-market valuation study of policies that have a visual impact.

5.6 Areas for future research

This study has shown that a VR treatment can improve the reliability of choice responses in terms of choice error and scope sensitivity. A better test of external validity of choice experiment results is to compare with actual choices made in the real world. An interesting avenue of research would be to design an experiment that tests the reliability of stated preferences versus real choices, with and without VR. For example, the choice of which site to visit.

Virtual environments developed using the tools I 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. There are many providers of pre-made models that can be downloaded (sometimes for free) and used within the Sketchup modelling software. The virtual environment can also include sound and animations such as a day/night cycle or crashing waves. Simple animations can be created in Sketchup. More research is required to confirm whether more realistic virtual environments outperform simple ones, and to what extent and under what circumstances the extra development effort is a worthwhile investment.

Another issue to consider is how best to incorporate attributes that cannot easily be visualised, such as water contaminants or cost. These attributes do not necessarily need to be presented in a traditional table. They could be embedded in the virtual environment, for example as a warning sign in front of a beach that is unsuitable for bathing.

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 actions are recorded. However, providing an interactive experience introduces additional technological and methodological complications. Controlling and recording what people look at requires more advanced web programming skills than simply embedding a video. Respondents with old, unsupported browsers or slow internet connections may not be able to participate.

5.7 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. doi:10.1016/S1366-5545(02)00047-9

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. doi: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. doi:10.1016/0148-2963(92)90058-j

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

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. doi:10.1007/s10640-008-9242-8

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. doi:10.1111/j.1467-8489.2009.00469.x

Campbell, D., & Erdem, S. (2014). Position bias in best-worst scaling surveys: a case study on trust in institutions.

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. doi: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. doi:10.1006/jeem.1997.0988

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

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). doi: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. doi:10.1006/jeem.2001.1199

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.

Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. (2010). The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity, *29*, 393–421. doi:10.1287/mksc.1090.0508

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. doi:10.1016/j.jeem.2008.08.002

Graeme, M., Dahm, J., & Kendal, H. (2010). *Coromandel Peninsula - Ecological Assessment of Natural Character*.

Greene, W. H., & Hensher, D. A. (2010). Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification. *Empirical Economics*, *39*(2), 413–426. doi:10.1007/s00181-009-0310-x

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.

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

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

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. doi: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. doi: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. doi:10.1016/j.jfe.2011.04.001

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.

Mitchell, R. C., & Carson, R. T. (1989). Using surveys to value public goods: the contingent valuation approach. *Resour. for the Future, Washington, DC*.

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

Sandor, Z., & Wedel, M. (2001). Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, *38*(4), 430–444.

Scarpa, R., Gilbride, T. J., Campbell, D., & Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics*, *36*, 151–174. doi:10.1093/erae/jbp012

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. doi:10.1093/ajae/aaq174

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. doi: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. doi:10.1023/A:1020262503119

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

Thomson, J. (2003). *Coastal values and beach use survey report*. Waikato Regional Council.

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

5.8 Appendix

5.8.1 Choice cards

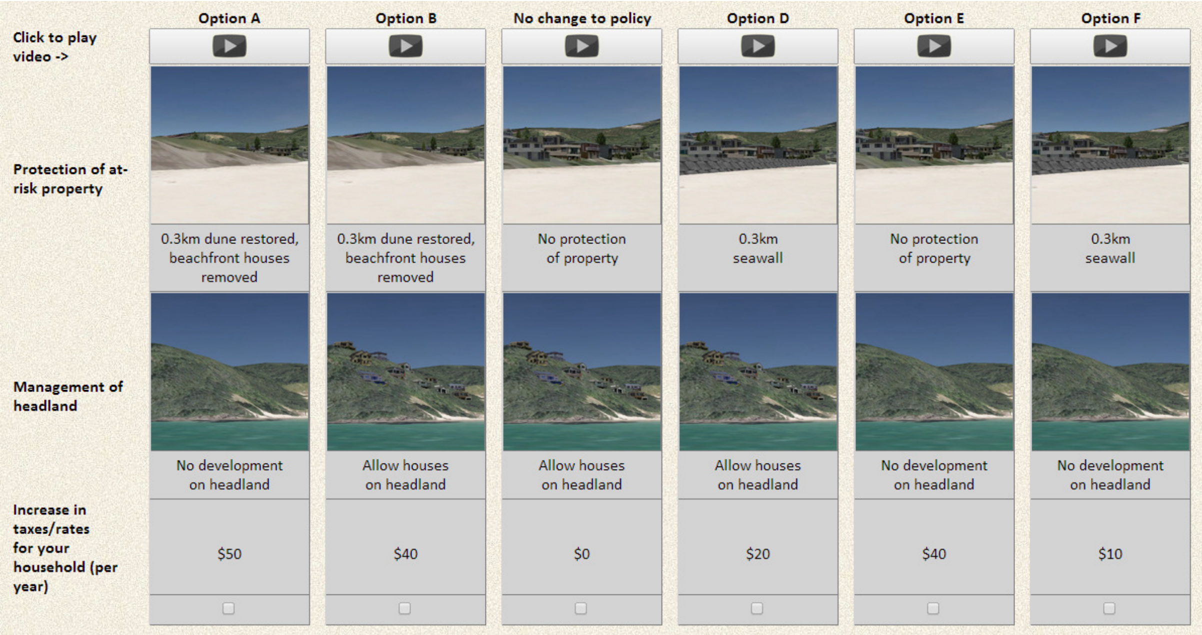

Figure 6 - Example beach choice card

	Current protection	Option A	Option B
Length of seawalls protecting high risk property *Currently 2.1km out of a total 14km at risk*	15%	15%	15%
Length of seawalls protecting medium risk property *Currently 0km out of a total 3.6km at risk*	0%	0%	100%
Length of seawalls protecting public infrastructure *Currently 4.3km out of a total 8.3km at risk*	50%	0%	100%
Increase in accommodation costs after 10 years *(per night)*	+ $30 *(20% increase)*	+ $100 *(66% increase)*	+ $75 *(50% increase)*

Figure 7 - Example whole-peninsula choice card