

On: 10 May 2013, At: 15:18

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Environmental Economics and Policy

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/teep20>

Constructed preference stability: a test-retest

Roy Brouwer^a

^a Department of Environmental Economics, Institute for Environmental Studies, Vrije Universiteit Amsterdam, De Boelelaan 1087, 1081, HV, The Netherlands

Published online: 18 Jan 2012.

To cite this article: Roy Brouwer (2012): Constructed preference stability: a test-retest, Journal of Environmental Economics and Policy, 1:1, 70-84

To link to this article: <http://dx.doi.org/10.1080/21606544.2011.644922>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Constructed preference stability: a test–retest

Roy Brouwer*

*Department of Environmental Economics, Institute for Environmental Studies, Vrije Universiteit
Amsterdam, De Boelelaan 1087, 1081 HV, The Netherlands*

(Received 11 October 2011; final version received 13 November 2011)

Preference construction has been argued to undermine the use of stated preference results like willingness to pay (WTP) values in policy evaluations. Preferences constructed at the time of the valuation task are expected to be unstable and hence unreliable. Choice repetition and experience have been shown to stabilize preferences and increase choice consistency. The objective of this article is twofold. First, to examine the effect of self-reported construction bias on stated preferences during a repeated choice contingent valuation (CV) study. Second, to test the effect of choice repetition on the stability of the articulated preferences in a test–retest experiment using a single and double bound CV referendum elicitation format. As expected, unfamiliarity with the environmental good in question results in a significant impact of the survey instrument on preference construction. This in turn influences the respondent decision to participate in the contingent market, but after controlling for self-selection, cognitive and motivational characteristics, not the size of the WTP value. Evidence is found that choice experience filters out possible construction bias. In the test–retest experiment, choice repetition furthermore generates more stable and consistent preferences compared to the results of the NOAA endorsed ‘one-shot’ single bound approach in CV.

Keywords: preference construction; preference stability; contingent valuation; test–retest; biodiversity; marine reserves; North Sea

1. Introduction

This article focuses on the effect of information provision and choice repetition in a contingent valuation (CV) study on preference construction and preference stability. Although there exists an impressive body of empirical evidence with respect to the influence of information in stated preference research (e.g. Ajzen *et al.* 1996, Munro and Hanley 1999, Berrens *et al.* 2004), the issue of preference construction and how to adequately address constructed preferences has been relatively under-investigated by environmental economists. Constructed preferences have been argued to undermine the validity and reliability of stated preferences (Spash 2002). On the other hand, choice experience has been shown to stabilize preferences and increase choice consistency (e.g. DeShazo and Fermo 2002). The ‘discovered preference hypothesis’ (DPH) postulates that preferences may be ill-defined when respondents come to a hypothetical market, but that a set of stable intrinsic values already exists

*Email: roy.brouwer@ivm.vu.nl

(Braga and Starmer 2005). Through repetition, individuals will gradually discover their exact preferences and adjust their behavior accordingly.

Bateman *et al.* (2008) find support for the DPH in a series of double bound CV questions for different goods. They argue that because of stabilizing learning effects it is the last response in a series of valuation questions rather than the first that yields the most reliable value statement, hence questioning the NOAA endorsed 'one-shot' single bound dichotomous choice (DC) approach in CV. In this article, the stability of stated preferences based on these two preference elicitation instruments, the single and double bound DC format, is tested in a test–retest experiment. At the same time, the effect of the preference elicitation instrument and information provided on preference construction is examined, based on respondent self-reported influence of the elicitation instrument, similar to existing approaches in the CV literature to assess respondent preference certainty (e.g. Alberini *et al.* 2003).

The main objective of this study is therefore twofold. First, to assess the effect of the preference elicitation instrument on preference construction in terms of individual willingness to pay (WTP) for a public environmental good. Second, to test the stability of these preferences in a test–retest experiment whilst accounting for choice repetition and experience. Respondents participating in a CV survey were contacted again after a couple of weeks and asked the same valuation questions again. In order to test the effect of choice experience on the stability of the originally articulated preferences, one group of respondents (the experimental group) was asked the double bound DC question, and another control group only the single bound DC WTP question. The hypothesis tested in this study is that preferences are more stable in time and hence more reliable for inclusion in policy evaluations such as cost-benefit analysis when based on a repeated double bound choice elicitation structure compared to a 'one-shot' single bound format. The overall CV survey format used in this study complies with guidelines to mitigate problematic preference construction when asking lay public to go through a novel and complex choice task (Payne *et al.* 1999).

The article is structured as follows. The issues surrounding preference construction and preference stability are further elaborated in Section 2. Section 3 presents the research design while results are presented in Section 4. Finally, Section 5 concludes.

2. Preference construction and stability

In micro-economic theory, it is assumed that individuals know their preferences and that these preferences are stable and coherent (Rabin 1998, Brown *et al.* 2008). Based on existing preferences, respondents in a CV survey are assumed to be able to compare, rank and trade-off the benefits of different levels of environmental good provision and increases in environmental quality against the costs involved in terms of personal money income they are being asked to give up. Underlying the economic concept of value is the assumption that individuals' preference orderings are stable and independent of irrelevant contextual aspects. However, lack of familiarity and experience with public goods and their valuation may undermine these *a priori* assumptions (e.g. Shaikh *et al.* 2007). Where individuals have little experience with a good, WTP will be determined under the uncertain condition in which they have to imagine good provision. They may rely on a number of factors to determine their values, one of which is the perceived cost of provision. The complexity of the

valuation scenario and the money value given in the standard DC WTP question may provide an implicit value cue and be used by respondents in a CV survey to arrive at their value for a good. Contrary to predicting choice behavior in other behavioral sciences such as marketing and decision theory, empirical evidence explicitly addressing preference construction in economic valuation research is rare (e.g. Slovic 1995, McFadden 1999, Spash 2002, Bateman *et al.* 2008).

Although preferences and expressions of preferences are very likely to be influenced by the information provided during the time of the choice task and the ability of respondents to fit and position this information (selectively) into their mental value system, this does not necessarily mean, as Payne *et al.* (1999) point out, that there is no 'true' value to be measured. Important to know is what exactly is measured and how the information provision shapes preferences. More specifically, Fischhoff *et al.* (1999) stipulate that constructed preferences can arise both when people hold well-articulated values and when they lack them. It is generally acknowledged in the psychology literature that choices and values are highly context dependent, and if more generally applicable, this implies that there is always – to some degree – an ongoing process of preference construction, where people construct preferences depending upon the options at hand (Tversky 1996). Equally, the more ambiguity in someone's preferences, for example due to a lack of familiarity with the good in question, the more expressed preferences will be subject to procedural and descriptive influences (Schkade and Payne 1994). It is this procedural (in)variance that is fundamental to theories of rational choice and raises questions about the nature of human values (Slovic 1995). The observed phenomenon of arbitrary coherent behavior (Ariely *et al.* 2003, 2006), where initially constructed *absolute* values based on a combination of both relevant and irrelevant contextual factors appear to be *relatively* coherent and consistent with economic theory, adds to the complexity of our limited understanding of these human values, especially in hypothetical settings for non-market goods like CV.

In order to avoid preference construction for public environmental goods, Fischhoff and Furby (1988) argue that a 'satisfactory transaction' in the traditional neo-classical sense, where people are fully informed, uncoerced and able to identify their own best interests, can only take place if the environmental good, the method of payment, and the constructed market are well defined and understood by the individual. An update and extension of the various components making up such a transaction is provided by Fischhoff *et al.* (1999).¹ Similarly, Cummings *et al.* (1986) argued in their first state-of-the-art report that the use of the CV method should be restricted to situations which best emulate consumer markets. Subjects should understand and be familiar with the commodity to be valued and they should have had (or be allowed to obtain) prior valuation and choice experience with respect to consumption levels of that commodity. Through repetition respondents are expected to be capable of making more precise and consistent decisions, because they learn about the survey format, the associated (hypothetical) market environment and their own preferences (List 2003). Hoefler and Ariely (1999) show that preference stability is indeed positively correlated with choice experience (single and repeated choice) and choice effort (easy choice versus hard choice). In the latter case, a higher level of effort leads to more stable preferences, but less preference strength, meaning that respondents facing a hard choice are less certain of their preferences than respondents facing an easy choice.

3. Research design

3.1. Testing for choice experience and preference stability

In order to be able to test for the effect of choice repetition and experience on the temporal stability of stated preferences, a single bound (SB) and double bound (DB) CV referendum model format was developed following Hanemann *et al.* (1991). The DB format is an extended version of the SB format (Equation (1)) where WTP depends on a variety of observable factors x_i , including bid price, and unobservable factors captured in the error term ε_i . In the DB format respondents are asked a second follow-up WTP question after the first WTP question (Equation (2)):

$$\text{WTP}_i^1 = x_i' \beta + \varepsilon_i \quad (1)$$

$$\text{WTP}_i^2 = (1 - \gamma) \text{WTP}_i^1 + \gamma B_1 + \delta \quad (2)$$

where γ is the parameter reflecting anchoring on the starting bid B_1 and δ is a shift parameter. If there is no starting point bias $\gamma = 0$. Likewise incentive incompatibility means that $\delta = 0$ (Aadland and Caplan 2004).

Hence, respondents are either asked one or two DC WTP questions in the control and experimental group respectively. Based on the existing literature (e.g. Cameron and Quiggin 1994, Alberini 1995), a priori expectations are that the SB elicitation format generates higher and more variable WTP values than the DB format. The DB elicitation format typically improves the statistical efficiency of referendum-based welfare measures. Possible reasons for the SB-DB gap are given in Carson and Grooves (2007) and relate primarily to incentive compatibility differences. Bateman *et al.* (2008) add to this the learning effect, arguing that '*the imprecise estimates in initial valuations appear inconsistent with a-priori well-formed preferences*' (ibid, p.131). Due to choice repetition and experience, the DB format is therefore expected to produce more stable preferences and WTP values in the test-retest experiment. Including the retest, where respondents in the experimental and control group are asked the same WTP question(s) again two weeks after they participated in the original CV study, respondents are asked either two or four WTP questions for one and the same good (at one or two bid price levels respectively at two different points in time). The survey design is detailed in the next section.

3.2. Survey design

Following the recommendations in Payne *et al.* (1999) to avoid problematic preference construction, the survey questionnaire was carefully developed during a pre-test procedure over a period of 4 weeks. Draft versions of the questionnaire focusing on protection of marine habitat and biodiversity in the Dutch North Sea were tested in two different rounds of pre-tests based on 70 face-to-face interviews at two different beach resorts along the Dutch coast. This led to modifications of the design, among others the presentation of information statements in short text boxes to avoid information overload and respondent fatigue.

A mail survey was used to give respondents time to carefully think through the valuation task at hand, that is, their WTP for the protection of marine biodiversity in the North Sea through the establishment of marine protected areas. In their own home, respondents were expected to be able to complete the questionnaire in their

own time without any of the potential biases associated with face-to-face interviews (e.g. Ethier *et al.* 2000). The mail questionnaire consisted of 45 questions, followed Dillman's (1978) 'total design method', and was sent out as a booklet. The questionnaire started off with a series of questions related to respondent beach visit behavior, perception, and concerns related to the current state of the North Sea.

Respondents were then informed about the actual state of the North Sea, existing and possible future threats and the development of new policy based on marine protected areas. Respondents were asked how familiar they are with the presented information and following Fischhoff *et al.* (1999) how much they trusted this information. After that, they were presented with three possible future policy and management scenarios for the North Sea: continuation of the current situation where nothing changes or a situation where ecologically vulnerable zones become designated marine protected areas, either fully protected where all economic activities are banned or where economic activities are allowed albeit conditional to certain restrictions and under strict control. It was emphasized that the protection of ecologically vulnerable zones where economic activities are restricted or banned involves a trade-off between economic and ecologic interests and hence comes at a cost. Following this information statement, respondents were asked a series of WTP questions. In order to avoid myopic behavior and stimulate respondents to think carefully about their preferences and WTP for the good in question, the valuation procedure consisted of a number of steps. First, respondents were asked if they would be willing to pay in principle for marine protected areas. This WTP question was context free. There is no reference to any payment vehicle or money amount. If respondents were not willing to pay in principle, they were asked for their motivation. Respondents who were willing to pay were also asked why. They were then asked the first bound DC WTP question where income taxation is used as the payment mode and payment frequency is annual over the next 10 years, followed by the second bound WTP question (except in the control group), and finally an open-ended maximum WTP question. The start bid ranged between 5 and 250 Euros per household per year.

Finally, at the end of the questionnaire, a number of manipulation checks were included to test the quality of the information respondents received and the influence this information had on their stated WTP. This included how easy or difficult it was to answer the WTP questions, how clear it was what they were being asked to pay for, and what the role was of the provided information in their value formulation and articulation. In this latter case, the self-reported influence of the preference elicitation instrument was measured on an itemized measurement scale from 0 (no influence at all) to 4 (a lot of influence), and was phrased as follows:

Compared to what you knew and thought before you received this questionnaire, how much did the questionnaire and the information in it influence the value you attach to the protection of the North Sea?

The results are presented in the next section.

4. Results

4.1. Sample characteristics

The mail survey was sent out to 7000 randomly selected households across the Netherlands in October 2006. The response rate was 17%, which is low, but not unusual for mail surveys. The sample is more or less representative for the

population from which it was drawn. Just over half of the respondents (51%) were males. The average age in the sample was 45 years. Respondents had to be 18 years or older to be able to participate in the survey and varied in age between 18 and 96 years. A third of the sample had a household with children, and 16% of the sample population was retired. A relatively high share of the sample population (54%) was higher educated (i.e. high school, college or university degree) and member of or donator to an environmental protection organization (42%), suggesting some degree of self-selection. The average donation level to environmental organizations was 75 euro per household per year. Net monthly household income was 2250 euro, which corresponded to the national average (Statistics Netherlands, 2008). Also the geographical distribution of the response was representative compared to the whole population. Almost half of the sample lived in a province along the Dutch coast.

A priori knowledge, awareness and concern about the valuation object were measured in different ways. Respondents were asked for their perception of sea water quality, their concern about coastal bathing water quality and the quality of fish caught for human consumption. A priori concern about the North Sea is used as another indicator of pre-existing value judgements and preferences for the valuation object in question. They were also asked on an itemized measurement scale from 0 (completely inadequate) to 4 (more than adequate) how informed they felt about sea water quality in the North Sea before participating in the survey. Note that using these indicators does not mean that respondents' concerns are correct and align with scientific findings or expert assessments. They were simply used here to assess whether or not respondents had prior judgements or preferences.

When asked whether the environmental state of the North Sea had improved or deteriorated over the past 10 years, as many as 43% of the sample ticked the 'I don't know' option. When asked for their perception of current seawater quality, 11% did not know. Thirteen percent believed the North Sea was polluted, 35% thought it was clean and 40% stated a neutral answer, that is, not clean, but also not polluted. Just over a quarter of all respondents felt they were being adequately informed about seawater quality in the North Sea by the responsible authorities.

The distribution of the respondents across the different categories of concern is presented in Figure 1. Three percent ticked the 'I don't know' option, while 15–25% of the sample was not in any way concerned about bathing water quality or fish consumption. These indicators have to be interpreted with the necessary care. If

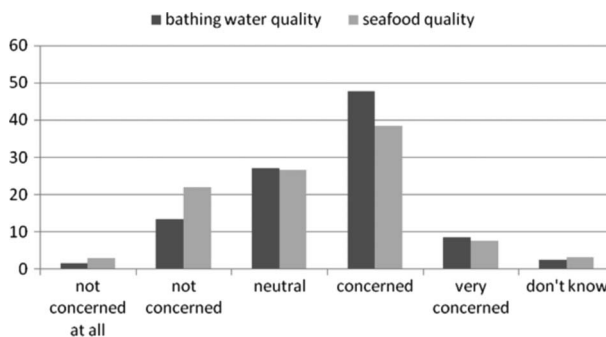


Figure 1. Frequency distribution (%) of public concern about North Sea bathing water and seafood quality.

someone was not concerned about the state of the environment, this does not necessarily mean that he or she does not have well-formed preferences for changes in the environment. A low level of concern is interpreted as a low level of what has been labeled 'product commitment' in marketing research (e.g. Martin and Goodell 1991). The implication here is that low levels of commitment reduce the ability of preference manipulation. A low, but statistically significant negative correlation (Spearman's $\rho = -0.24$; $p < 0.001$; $n = 1018$) exists between respondent concern about bathing water quality and how informed someone felt about seawater quality. Respondents who were less concerned were more comfortable with current levels of information provision.

Examining the degree of respondent familiarity with the information provided in the questionnaire, 6% of the sample population said they never heard of this information before, while 25% indicated not to be familiar with the information. So just over 30% of the sample were unfamiliar with the provided information. Fifty-one percent had heard about what they were told in the questionnaire before, but only 16% felt familiar and 3% very familiar with the information.

Of interest here is also to what extent respondents actually believed and trusted what they were being told. Here, a distinction was made between the information provided about the current state and threats to the North Sea and the future policy scenarios. A majority of the sample trusted the presented information. Three percent had no trust in the information provided about the current situation, while 15% did not have any confidence in the presented future policy scenarios. Less than 2% of the sample ticked 'I don't know'. A significant positive correlation is found between trust in the baseline and policy scenario description (Spearman's $\rho = 0.26$; $p < 0.001$; $n = 1123$), but not between respondent familiarity with the information provided in the survey and trust and confidence in the policy scenarios.

Respondents were also asked whether the information provided in the questionnaire sufficed to help them choose between the alternative policy options to protect the North Sea and determine their level of financial commitment for their most preferred option through WTP.² Three percent were not sure, while 80% said it was sufficient or more than sufficient. Seventeen percent felt the information was insufficient to answer the WTP question. As expected, a significant positive correlation exists between how respondents evaluated the level of information provision to answer the WTP questions and how clear it was what they were being asked to pay for (Spearman's $\rho = 0.37$; $p < 0.001$; $n = 1101$). Those who felt the information was insufficient struggled answering the WTP question. Respondents were asked to indicate on an itemized measurement scale from 0 (not difficult at all) to 4 (very difficult) how difficult it was for them to answer the WTP question. This variable correlates significantly with the adequacy of information provision too (Spearman's $\rho = -0.22$; $p < 0.001$; $n = 1100$). Moreover, the more difficult it was to state WTP, the less clear it was what they were being asked to pay for (Spearman's $\rho = -0.26$; $p < 0.001$; $n = 1138$). Eighty-five percent was clear or very clear what they were being asked to pay for. Only one percent did not know, and less than 10% (6.3%) were unclear about what they were paying for.

4.2. The effect of the preference elicitation instrument on preference construction

As expected, a significant negative correlation exists between how familiar respondents were with the information provided in the questionnaire and the

self-reported influence this had on their value judgment (Spearman's $\rho = -0.22$; $p < 0.001$; $n = 1155$). The less familiar, the more the information influenced the value someone attached to the protection of the North Sea. Most respondents (41%) indicated that the survey had some degree of influence on their value judgment, while almost a third claimed it had no effect. Five percent said it had a lot of effect. Self-reported instrument influence or construction bias was linked to both the referendum WTP and open-ended maximum WTP.

In the case of the referendum DC WTP values, two categories were created to ensure there are sufficient observations per category: (i) no influence *versus* (ii) some to a lot of influence. The single and double bound WTP values across these two main categories for the experimental group are presented in Table 1. A significant difference exists at the 10% level between the first single bound WTP values for respondents who claimed the preference elicitation instrument had no influence on their value statement and respondents who said it had some or a lot of influence. However, this difference is not statistically significant for the repeated choice double bound WTP values,³ suggesting that the self-reported construction bias disappears as a result of choice experience. How choice repetition and experience affect the temporal stability of stated WTP values will be presented in the next section. Before turning to these results, the open-ended WTP results are presented here too to enable comparison with the approach and results found in Spash (2002).

The distribution of respondents across the original self-reported instrument influence categories and associated mean maximum WTP is presented in Figure 2. The differences between the mean WTP values in Figure 2 are statistically significant at the one percent confidence level (Kruskal–Wallis (KW) $\chi^2 = 14.207$; $p < 0.007$), suggesting as before for the single bound WTP values that preference construction during the survey significantly influenced stated WTP.⁴ This is especially the case for respondents who said that the survey had a lot of influence on their value judgment. Mean WTP is in this case significantly higher compared to respondents who said the survey had no influence on their value judgment.

Despite extensive pre-testing, a substantial share of the sample population protested against the contingent market (13%).⁵ Common practice in CV research is to omit these protest responses from the analysis (Dziegielewska and Mendelsohn 2007). However, removal of protest response from further analysis of the WTP bids assumes that they are randomly distributed across protest and non-protest respondents, not biasing the sample representativeness and the estimated WTP

Table 1. The effect of self-reported construction bias on the DC WTP results.

	Single bound		Double bound	
	No self-reported bias	Self-reported construction bias	No self-reported bias	Self-reported construction bias
Mean WTP	65.3	92.5	92.3	97.9
Standard error	12.1	9.4	5.9	6.0
<i>N</i>	376	342	293	289
Difference (%)	29.4		6.0	
<i>t</i> -test statistic	1.775		0.661	
<i>p</i> -value	0.076		0.509	

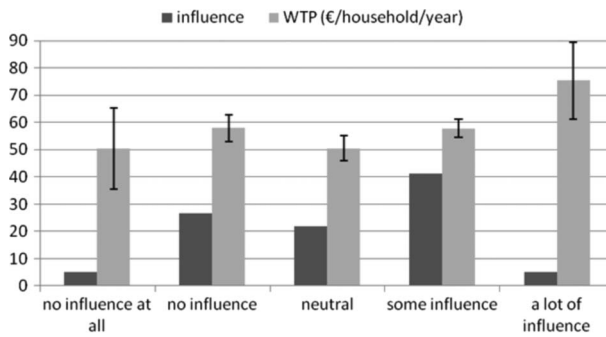


Figure 2. Frequency distribution (%) of the self-reported construction bias on stated open-ended WTP and standard errors.

value parameters. In order to avoid potential selection bias, the Heckman two-stage estimation procedure (Heckman 1979) is applied, where two functions are estimated sequentially, one specifying the decision to participate in the contingent market and one specifying the open ended WTP bid function. This procedure tests whether participation is random between the two sub-samples (of protesters and non-protesters). If nonrandom, the factors that determine participation in the contingent market may be correlated to the factors that determine the WTP values, resulting in a violation of the assumption of identically independently distributed errors. The WTP value is in that case conditional on the decision to participate in the contingent market.

The Heckman two-stage procedure accounts for this correlation through the inclusion of the inverse Mills ratio, estimated during the first stage in a probit model and representing the probability of a respondent participating in the contingent market. If the inverse Mills ratio is not significantly different from zero, the null hypothesis of no selection bias cannot be rejected, suggesting that the estimated second-stage model is representative for the whole sample, including those respondents who refused to pay. The results of the two-stage estimation procedure are presented in Table 2. Only variables that were found to be significant at least at the 10% confidence level are presented. Variable definitions are presented in Table 3.

A significant negative correlation is found between the error terms in the (first stage) selection and (second stage) primary model. The negative sign for the lambda (λ) suggests that not accounting for selection bias underestimates the probability of a positive WTP response. Hence, the factors that make participation more likely tend to be associated with higher WTP. These factors include respondent education level, membership of an environmental organization, whether the respondent feels any responsibility for the current state of the North Sea and the self-reported influence of the preference elicitation instrument on the value a respondent attaches to the protection of the North Sea. In the latter case, the positive sign of the coefficient estimate indicates that the more the survey instrument influenced the respondent's value judgment, the more likely the person was to participate in the contingent market. However, this variable did not have a significant impact on the actual bid amount in the primary model. So, whilst controlling for a variety of explanatory factors in the multivariate regression analysis, self-reported preference construction only plays a role in the decision (whether or not) to participate in the contingent market, not in relation to the size of the bid amount as the univariate KW test results

Table 2. Heckman two-stage multivariate WTP regression results.

	Variable summary				Model coefficient estimates		
	Mean	Std	Min	Max	Estimate	Ste	$p \leq$
<i>Selection model</i>							
Higher education	0.54	0.50	0	1	0.645	0.100	0.001
Instrument influence	2.14	1.03	0	4	0.166	0.038	0.001
Sense of responsibility	0.55	0.50	0	1	0.220	0.105	0.036
Member	0.42	0.49	0	1	0.256	0.107	0.017
λ (inverse Mill's ratio)					-0.538	0.232	0.021
<i>Primary model</i>							
Constant					-0.525	0.549	0.339
Start bid	83.90	78.59	5	250	0.476	0.025	0.001
Household income	2256	1358	750	7500	0.250	0.062	0.001
Gender	0.49	0.50	0	1	-0.162	0.066	0.014
Most preferred option	0.38	0.49	0	1	0.135	0.067	0.043
Concern seawater quality	2.49	0.88	0	4	0.150	0.039	0.001
Main motivation	0.68	0.46	0	1	0.222	0.070	0.002
Familiarity	1.86	0.85	0	4	0.075	0.039	0.056
Trust	2.76	1.02	0	4	0.059	0.033	0.078
<i>Model summary</i>							
Wald chi-square					420.45		0.001
N					759		

Note: Std, standard deviation; Ste, standard error.

Table 3. Variable description.

Variable	Type	Description
Higher education	Dummy	Variable has the value 1 if respondent completed high school, college or university degree
Instrument influence	Linear	Level of influence elicitation instrument on public good value judgment (protection of North Sea)
Sense of responsibility	Dummy	Variable has the value 1 if respondent feels in any way responsible for the current state of the North Sea
Member	Dummy	Variable has the value 1 if respondent is a member of or donator to an environmental organization
Household income	Linear	Respondent monthly net income
Gender	Dummy	Variable has the value 1 if respondent is female
Most preferred option	Dummy	Variable has the value 1 if most preferred option is a fully protected marine protected area without any economic use
Concern seawater quality	Linear	Level of concern about seawater quality
Main motivation	Dummy	Variable has the value 1 if nonuse is the dominant motivation for WTP
Familiarity information	Linear	Familiarity with the information provided in the questionnaire
Trust in valuation scenario	Linear	Level of trust in the valuation scenario

associated with Figure 2 suggested. Higher educated respondents, members or donators to an environmental organization and respondents who feel a responsibility towards the North Sea are also more likely to participate.

In the primary model, the start bid and household income influence the maximum amount of money respondents were willing to pay in a positive way. The start bid provides an important value cue to respondents for their open-ended WTP, and is therefore an important source of procedural variance. Respondents with higher incomes are, as expected, willing to pay significantly more than respondents with lower incomes. Moreover, women are willing to pay less than men for the protection of the North Sea. Other important determinants of maximum WTP are respondent preferences for the level of protection of marine protected areas (those who prefer completely closed off areas are willing to pay more than those who prefer protected areas with limited economic activities) and motivational orientation. The presence of nonuse values significantly increases WTP. The nonuse dummy variable has the value 1 if the respondent indicated that nonuse values, measured as 'the most important reason I am willing to pay is related to the value I attach to the environment and future generations', dominate over and above use values, measured as 'the most important reason I am willing to pay is related to the value I attach to my own health and the quality of bathing water and fish food'. Finally, as expected, respondent concern for bathing water quality, familiarity with the information provided in the questionnaire and trust in the policy (valuation) scenario also influence WTP in a positive way. A formal test of the effect of choice repetition on the stability of stated preferences will be presented in the next section.

4.3. *The effect of repeated choice on preference stability*

The stability of stated preferences was tested through a telephone follow-up survey, targeting those respondents who agreed to the telephone interview by including their telephone number in the original mail survey upon return of the questionnaire (30% of the sample). Before answering the same referendum WTP questions again, respondents were first asked whether they still remembered their original WTP replies in the mail questionnaire. The share of respondents who did not remember their original WTP in the experimental group increases from 5% for the first WTP question (in principle) to 10% for the second (start bid), 13% for the third (follow-up bid) and almost 40% for the fourth and last WTP question (open ended maximum WTP). Vice versa, the number of respondents who correctly remembered their exact original answer decreases from 92% for the first WTP question to 45% for the fourth and last WTP question. So, it is easier for respondents to remember whether or not they were willing to pay in principle than a specific bid amount or their maximum WTP amount.

The stability of stated preferences is examined again through the estimation of standard bivariate probit models and corresponding WTP welfare estimates (e.g. Cameron and James 1987) for the experimental and control group. The results are presented in Table 4. The impact of choice repetition on the stability of stated preferences is tested based on the observed difference between the test and retest results for the single and double bound elicitation format (last row in Table 4). Generally, the retest yields smaller WTP values than those originally stated. So, respondents tend to modify their original replies downwards when given time to think, although the differences are not significantly different.⁶ The variation in stated WTP in the test and retest can be measured through the variation coefficient (standard error divided by mean WTP) and is also more or less the same.⁷

Table 4. Test–retest WTP results and stability of stated preferences.

	Single bound, control Group		Single bound, experimental Group		Double bound, experimental Group	
	Test (t_0)	Retest (t_1)	Test (t_0)	Retest (t_1)	Test (t_0)	Retest (t_1)
Mean WTP	186.5	162.9	142.0	137.4	109.2	107.5
Standard error	33.0	25.6	17.6	18.0	9.0	8.7
<i>N</i>	89	89	181	181	177	177
Change (%)		-12.7		-3.2		-1.6

The first bound WTP values for the control group are based on the one-shot value elicitation procedure. Respondents are asked in the retest the same one-shot WTP question again after they have been given time to think about their preferences for the public environmental good involved. The first bound WTP values for the experimental group are based on a double bound repeated choice procedure. Respondents have more choice experience in this experimental group and this is expected to have a stabilizing effect on stated preferences in the retest. This a priori expectation is confirmed when examining the difference between the test and retest WTP values. Although the difference between the single bound test and retest results are not significantly different based on a simple *t*-test for either the control and the experimental group,⁸ their relative difference is four times larger in the case of the control group than for the experimental group. The difference between the test and retest double bound WTP values is even smaller. As expected, the double bound WTP values are smaller and less variable than the single bound WTP values in the experimental group.

5. Conclusions

The main objective of the study presented in this article was twofold: to find empirical evidence of preference construction in CV research and investigate the temporal stability of these preferences in a test–retest experiment. Preference construction was measured by asking respondents after they went through a series of WTP questions to what extent the preference elicitation instrument influenced the value they attach to the public good they were being asked to value, referred to here as construction bias. Preference stability was measured through a within-subject experimental treatment by asking the same WTP questions two weeks later whilst accounting for choice experience in a single and double bound referendum elicitation format. Cognitive and motivational manipulation checks were included related to respondent a priori knowledge and awareness, attitudes and ethical stance towards paying for the public environmental good, familiarity with and trust in the information provided in the questionnaire and thinking tools like how much they pay already towards environmental protection generally as donators or members of environmental organizations.

Evidence of preference construction is found based on the self-reported influence the preference elicitation instrument had on value judgments. Almost half of all the respondents reported some degree of influence of the survey instrument and information provided. This effect was, as expected, strongest for respondents who

were least familiar with the information provided. Those who indicated that the survey had much influence on their value judgment stated a significantly higher WTP. However, whilst controlling for self-selection and the above mentioned cognitive and motivational variables, this effect disappeared and the survey instrument only appeared to influence the decision to participate in the contingent market, not the actual WTP amount. This is a subtle, but important difference compared to the results presented in Spash (2002), who finds a significant positive effect on the WTP amount of respondents who said that the information they were given changed their preferences. So, once control is included, for example regarding the level of concern about and familiarity with the public good in question, the influence of the elicitation instrument and information provision on the actual bid amount is not significant. Preference construction was limited to the decision to participate in the contingent market not the decision how much respondents were willing to pay.

Moreover, evidence is found that choice repetition and experience filters out possible construction bias. Although self-reported construction bias significantly affects the 'one-shot' single bound WTP values, this effect disappears in the case of the double bound WTP values. The stability of these stated preferences was investigated further through a test-retest experiment. The hypothesis tested here was the a priori expectation that choice experience has a stabilizing effect on constructed preferences and increases choice consistency. Respondents who participated in the original CV study were interviewed again two weeks later by phone and asked the same WTP questions again. One sub-sample of respondents was presented the NOAA endorsed single bound WTP elicitation format in the original CV study and another sub-sample a repeated choice double bound WTP format. Although the elicited values remain largely stable and invariant after being constructed, i.e. consistent with the initially articulated preferences and values, choice experience gained through the double bound elicitation format yields more stable and consistent preferences than the NOAA endorsed single bound approach in CV.

Acknowledgements

The research presented here was funded by the National Institute for Coastal and Marine Management (RIKZ) in The Hague, the Netherlands. The useful comments and suggestions from Stavros Georgiou, Anthony Ryan and Joop de Boer on a draft version of the paper are gratefully acknowledged. The usual disclaimer applies.

Notes

1. In those cases where respondents face a novel and relatively complex choice task, Payne *et al.* (1999) developed a 'building code' for 'good' preference construction. Problematic preference construction may be due, among others, to myopic decision frames, inappropriate problem representation, lack of comprehension of the information provided and avoidance of tradeoffs.
2. Most respondents preferred marine protected areas where some economic activities are allowed under strict conditions (58%), followed by zoning of areas where economic activities are banned altogether (38%). Three percent preferred the current situation. Only 1.5 percent of the sample population was unable to make up their mind and did not know which valuation scenario to choose.
3. The standard errors for the double bounded WTP values are estimated based on the Krinsky and Robb (1986) procedure.
4. This result is confirmed based on the Median test (chi square = 8.808) albeit at a lower confidence level ($p < 0.066$).

5. The most important protest reason was that respondents felt ‘the North Sea should be protected by law, not by asking people to pay for it’ (11%), followed by ‘the polluter should pay’ (1%), ‘protection of the North Sea is a government task’ (0.3%) and ‘lack of sufficient information’ (0.2%).
6. Test results available from the author.
7. The standard errors for the double bounded WTP values are estimated again based on the Krinsky and Robb (1986) procedure.
8. The outcome of the t -test is $t = 0.571$ ($p = 0.569$) for the control group and $t = 0.181$ ($p = 0.857$) in the case of the experimental group.

References

- Aadland, D. and Caplan, A.J., 2004. Incentive incompatibility and starting point bias in iterative valuation questions: comment. *Land Economics*, 80 (2), 312–315.
- Ajzen, I., Brown, T.C., and Rosenthal, L.H., 1996. Information bias in contingent valuation: effects of personal relevance, quality of information and motivational orientation. *Journal of Environmental Economics and Management*, 30 (1), 43–57.
- Alberini, A., 1995. Efficiency vs bias of willingness-to-pay estimates: bivariate and interval-data models. *Journal of Environmental Economics and Management*, 29, 169–180.
- Alberini, A., Boyle, K., and Welsh, M., 2003. Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty. *Journal of Environmental Economics and Management*, 45, 40–62.
- Ariely, D., Loewenstein, G., and Prelec, D., 2003. “Coherent arbitrariness”: stable demand curves without stable preferences. *Quarterly Journal of Economics*, 118 (1), 73–105.
- Ariely, D., Loewenstein, G., and Prelec, D., 2006. Tom Sawyer and the construction of value. *Journal of Economic Behavior & Organization*, 60, 1–10.
- Bateman, I.J., *et al.*, 2008. Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness. *Journal of Environmental Economics and Management*, 55, 127–141.
- Berrens, R.P., *et al.*, 2004. Information and effort in contingent valuation surveys: application to global climate change using national internet samples. *Journal of Environmental Economics and Management*, 47 (2), 331–363.
- Braga, J. and Starmer, C., 2005. Preference anomalies, preference elicitation and the discovered preference hypothesis. *Environmental and Resource Economics*, 32, 55–89.
- Brown, T.C., *et al.*, 2008. Reliability of individual valuations of public and private goods: choice consistency, response time, and preference refinement. *Journal of Public Economics*, 92, 1595–1606.
- Cameron, T.A. and James, M.D., 1987. Efficient estimation methods for closed-ended contingent valuation survey data. *Review of Economics and Statistics*, 69, 269–276.
- Cameron, T.A. and Quiggin, J., 1994. Estimation using contingent valuation data from a ‘dichotomous choice with follow-up’ questionnaire. *Journal of Environmental Economics and Management*, 27, 218–234.
- Carson, R.T. and Groves, T., 2007. Incentive and informational properties of preference questions. *Environmental and Resource Economics*, 37(1), 181–210.
- Cummings, R.G., Brookshire, D.S., and Schulze, W.D., 1986. *Valuing environmental goods: an assessment of the contingent valuation method*. Totowa: Rowman and Allanheld.
- DeShazo, J.R. and 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.
- Dillman, D.A. 1978. *Mail and telephone surveys: the total design method*. New York: John Wiley and Sons.
- Dziegielewska, D.A. and Mendelsohn, R., 2007. Does “No” mean “No”? A protest methodology. *Environmental and Resource Economics*, 38, 71–87.
- Ethier, R.G., *et al.*, 2000. A comparison of hypothetical phone and mail contingent valuation responses for green-pricing electricity programs. *Land Economics*, 76 (1), 54–67.
- Fischhoff, B. and Furby, L., 1988. Measuring values: a conceptual framework for interpreting transactions with special reference to contingent valuation of visibility. *Journal of Risk and Uncertainty*, 1, 147–184.

- Fischhoff, B., Welch, N., and Frederick, S., 1999. Construal processes in preference assessment. *Journal of Risk and Uncertainty*, 19 (1–3), 139–164.
- Hanemann, W.M., Loomis, J.B., and Kanninen, B., 1991. Statistical efficiency of double bounded dichotomous choice contingent valuation. *American Journal of Agricultural Economics*, 73, 1255–1263.
- Heckman, J.J., 1979. Sample selection bias as a specification error. *Econometrica*, 47, 153–161.
- Hoeffler, S. and Ariely, D., 1999. Constructing stable preferences: a look into dimensions of experience and their impact on preference stability. *Journal of Consumer Psychology*, 8 (2), 113–139.
- Krinsky, I. and Robb, A.L., 1986. On approximating the statistical properties of elasticities. *Review of Economic and Statistics*, 68, 715–719.
- List, J.A., 2003. Does market experience eliminate market anomalies? *Quarterly Journal of Economics*, 118, 41–71.
- Martin, C.L. and Goodell, P.W., 1991. Historical, descriptive and strategic perspectives on the construct of product commitment. *European Journal of Marketing*, 25 (1), 53–60.
- McFadden, D., 1999. Rationality for economists. *Journal of Risk and Uncertainty*, 19, 73–106.
- Munro, A. and Hanley, N.D., 1999. Information, uncertainty and contingent valuation. In: I.J. Bateman and K.G. Willis, eds. *Valuing environmental preferences: theory and practice of the contingent valuation method in the US, EU, and developing countries*. New York: Oxford University Press, 258–280.
- Payne, J.W., Bettman, J.R., and Schkade, D.R., 1999. Measuring constructed preferences: towards a building code. *Journal of Risk and Uncertainty*, 19(1–3), 243–270.
- Rabin, M., 1998. Psychology and economics. *Journal of Economic Literature*, 36, 11–46.
- Schkade, D.A. and Payne, J.W., 1994. How people respond to contingent valuation questions: a verbal protocol analysis of willingness to pay for an environmental regulation. *Journal of Environmental Economics and Management*, 26, 88–109.
- Shaik, S.L., Sun, L., and van Kooten, G.C., 2007. Treating respondent uncertainty in contingent valuation: a comparison of empirical treatments. *Ecological Economics*, 62, 115–125.
- Slovic, P., 1995. The construction of preference. *American Psychologist*, 50, 364–371.
- Spash, C.L., 2002. Informing and forming preferences in environmental valuation: coral reef biodiversity. *Journal of Economic Psychology*, 23, 665–687.
- Tversky, A., 1996. Contrasting rational and psychological principles in choice. In: R.J. Zeckhauser, R.L. Keeney and J.K. Sebenius, eds. *Wise choices: decisions, games and negotiations*. Boston: Harvard Business School Press, 5–21.