Towards a more structured selection process for attributes and levels in choice experiments: A study in a Belgian protected area

Peer-reviewed author version

JEANLOZ, Sarah; LIZIN, Sebastien; BEENAERTS, Natalie; Brouwer, Roy; VAN PASSEL, Steven & WITTERS, Nele (2016) Towards a more structured selection process for attributes and levels in choice experiments: A study in a Belgian protected area. In: Ecosystem Services, 18, p. 45-57.

DOI: 10.1016/j.ecoser.2016.01.006
Handle: http://hdl.handle.net/1942/21168

REVISED MANUSCRIPT

Towards a more structured selection process for attributes and levels in choice experiments: a study in a Belgian protected area

Jeanloz Sarah[a], Lizin Sebastien[a], Beenaerts Natalie[a], Brouwer Roy[b], Van Passel Steven[a,c], Witters Nele[a*]

[a]*Centre for Environmental Sciences (CMK), Hasselt University, Agoralaan, Building D, 3590 Diepenbeek, Belgium*

Email Sarah Jeanloz: sarah.jeanloz@uhasselt.be

Email Sebastien Lizin: sebastien.lizin@uhasselt.be

Email Natalie Beenaerts: natalie.beenaerts@uhasselt.be

Email Steven Van Passel: steven.vanpassel@uhasselt.be

Email Nele Witters: nele.witters@uhasselt.be

[b]*University of Waterloo, Department of Economics, 200 University Avenue West, Waterloo, ON, N2L 3G1, Canada*

Email Roy Brouwer: rbrouwer@uwaterloo.ca

[c]*University of Antwerp, Department Engineering Management, Prinsstraat 13, 2000 Antwerp*

Email Steven Van Passel: steven.vanpassel@uantwerpen.be

*Corresponding author:
Dr. Nele Witters
Postal address: Agoralaan, Building D, 3590 Diepenbeek, Belgium
E-mail address: nele.witters@uhasselt.be
Telephone number: +32 11 26 8748

ABSTRACT

The process of selecting attributes for inclusion in choice experiments frequently involves qualitative methods such as focus groups and interviews. In order for a choice experiment to be successful and the results to be valid, this qualitative selection process is essential. It often lacks rigour and is poorly described, particularly in environmental choice experiments. We propose a meticulous attribute and attribute-level selection process consisting of a scoring exercise and an interactive discussion. This paper provides a case study describing how attributes and attribute-levels were identified and selected for the National Park Hoge Kempen in Belgium. We carried out four focus groups and thirteen semi-structured interviews with various park stakeholders to select attributes from six categories: the four categories of ecosystem services (supporting, provisioning, regulating, cultural), infrastructure, and land use types. The top-ranked characteristics were nature conservation, natural forests, biodiversity refuge, wetlands, landscape variety, heathlands, air purification, and education. Both the scoring exercise and the interactive discussion contributed to the attributes selected for the CE. Following these, an ultimate expert consultation stage is recommended to approve both the attribute and attribute-level selection. The semi-qualitative protocol proposed in this paper can help practitioners and demonstrates how the results guide choice experiment design.

Key words: semi-qualitative protocol; scoring exercise; picture-based; 'satisfaction' question; ecosystem services

1. INTRODUCTION

Stated preference surveys have proven to be versatile valuation techniques for estimating both use and non-use values (Bennett and Blamey, 2001, Bateman, 2002, Rolfe and Windle, 2015). Choice experiments (CE), in particular, have been increasingly used in the ecosystem service (ES) and biodiversity domain to elicit public and stakeholder preferences for management interventions and policy changes (Birol and Koundouri, 2008). Survey respondents are presented with several choice tasks consisting of hypothetical alternatives (scenarios) framing an environmental good or service to be valued. These alternatives are composed of a number of attributes and attribute-levels. Neoclassical economists state that by trading off attribute levels and choosing the preferred alternative, respondents are assumed to maximize their utility while indirectly expressing their willingness-to-pay (McFadden, 1974). Classical and ecological economists have a more social constructivist perception of value formation, behaviour and choice (Vatn, 2009). They disagree with the utilitarian conception of values and argue that monetary valuation of public goods (*e.g.* biodiversity) fosters social inequality, focuses exclusively on individual preferences and ignores non-economic cultural values (Spash, 2002, Wilson and Howarth, 2002, Krasny *et al.*, 2014). Moreover, there is disagreement regarding the use of monetary valuation to elicit non-material values (Chan *et al.*, 2012).

Across research fields that apply CE, such as health care, marketing, transportation and environmental economics, the attribute generation process consists of two initial steps: 1) to identify policy alternatives and relevant attributes, and 2) to assign relevant attribute-levels. Attributes influence an individual's decision, thus ignoring relevant attributes in a CE biases findings (Lancsar and Louviere, 2006, Coast *et al.*, 2012). Stated preference approaches should be user-useful. In an ES context for instance, it is required that practitioners respond to stakeholder needs from the start and collaborate to achieve the protection of ES and guarantee the flow of these ES to beneficiaries (Cowling *et al.*, 2008). A sound attribute selection process, that entails both detailed reporting and rigorous application of qualitative methods, can reduce the complexity of choice tasks and therefore the cognitive burden associated with CE (Rolfe *et al.*, 2004). The latter issues may arise when respondents are asked to trade-off between multifaceted and unfamiliar goods and services such as those generally involved in environmental valuation (Hoyos, 2010). The initial stages of any stated preference valuation study has to be grounded on some kind of social elicitation process in order to inform environmental or other public policy decision-making (Brouwer *et al.*, 1999). These stages are essential if the problem of stakeholder unfamiliarity, that might occur when using stated

preference valuation methods, is to be surmounted (Hein *et al.*, 2006, Barkmann *et al.*, 2008, Cowling *et al.*, 2008).

Recent papers in health economics call for detailed reporting on the process of attribute generation for CE and argue that qualitative studies are best suited to derive attributes, since they reflect the perspective and experiences of the potential beneficiaries (Coast and Horrocks, 2007, Ryan *et al.*, 2009, Coast *et al.*, 2012, Kløjgaard *et al.*, 2012, Hiligsmann *et al.*, 2013, Abiiro *et al.*, 2014, Michaels-Igbokwe *et al.*, 2014). A list of possible attributes can be generated a priori from the literature, but this list must be upgraded through participative processes, such as focus groups, expert consultations and pilot testing. For attribute identification and attribute-levels assignment a wide variety of qualitative approaches is typically used, due to their suitability to identify attributes for CE (Bateman, 2002, Coast *et al.*, 2012, Kløjgaard *et al.*, 2012). Qualitative research methods include literature reviews, visits to the study area, exploratory surveys, expert and key informant consultations, focus groups and interviews (Bateman, 2002, Blamey *et al.*, 2002, Coast *et al.*, 2012, Abiiro *et al.*, 2014). Brouwer *et al.* (1999) demonstrated that respondents in a stated-preference survey favoured participatory approaches to inform environmental decision-making process.

In environmental CE, attributes may represent land use types (Hoyos, 2010, Shoyama *et al.*, 2013), ES (Barkmann *et al.*, 2008), biodiversity features such as plant and animal species (Cerda *et al.*, 2013), tourism facilities and activities (Chaminuka *et al.*, 2012), and geographical attributes such as location and size (Rolfe *et al.*, 2000). Environmental CE studies that have applied focus groups and or interviews to select attributes generally combine them with methods such as expert consultation, discussions, and literature research. However, where these two qualitative methods are applied for attribute generation, very often little or no description is provided, thus leaving room for doubt whether these are indeed all relevant and encompassing (Coast *et al.*, 2012, Abiiro *et al.*, 2014, Armatas *et al.*, 2014). Environmental CE studies which do not perform qualitative work assume that selecting attributes based on previous work, literature review or “discussions” suffices (Li *et al.*, 2004, Rajmis *et al.*, 2009, Liu and Wirtz, 2010). Information about the amount of time taken to select attributes and the type of stakeholders are frequently lacking. In the environmental economics domain, we are only aware of Armatas *et al.* (2014) who documented a detailed attribute selection process. They applied the Q-methodology, a non-monetary preference elicitation technique that can highlight ES that are suitable for valuation and salient to a wide range of stakeholders (Kløjgaard *et al.*, 2012, Armatas *et al.*, 2014).

This paper responds to and builds on the aforementioned health and ecological economics studies. We contribute to the need for more rigorous attribute selection processes in the environmental economics domain. Here, we propose an attribute selection process that is based on the most frequently used qualitative methods, *i.e*. focus groups (FGs) and semi-structured personal interviews (INTs). The participation of park stakeholders is necessary to select attributes that are relevant to them (demand-relevant) and that they would like to see change. This study provides an easy-to-use and transferable approach, considered as semi-qualitative, to support the selection of attributes for environmental CE. Our final CE will aim to understand preferences of and trade-offs made by visitors for the characteristics (*i.e*. future CE attributes) of the National Park Hoge Kempen in Belgium.

The remainder of this paper is organized as follows. In Section 2, we describe our case study, the National Park Hoge Kempen in Belgium. In Section 3, we outline the rationale for the research methods chosen, and propose a framework for the identification and selection of CE attributes and attribute-levels. Then, we thoroughly outline our approach in five successive stages, including the methodology (Section 4: stage 1 to 3) and the results (Section 5: stage 4 and 5). In Section 6, we discuss the results and the protocol's shortcomings, while Section 7 holds the conclusion and provides general recommendations.

2. CASE STUDY: THE NATIONAL PARK HOGE KEMPEN

The study focuses on the National Park Hoge Kempen (NPHK), located in the Province of Limburg in the East of Belgium (Fig. 1). The NPHK (inaugurated in 2006) is surrounded by six municipalities with a total of about 163,500 inhabitants, equivalent to a population density of 450/km^2 (average density in Flanders 539/km^2). This first and only Belgian national park covers an area of approximately 6000 ha with a rich variety of habitats, including heathlands. This cultural North-West-European landscape, rich in biodiversity, has experienced a drastic surface reduction in the past decades due to urbanisation and tree planting for the coal mining industry.

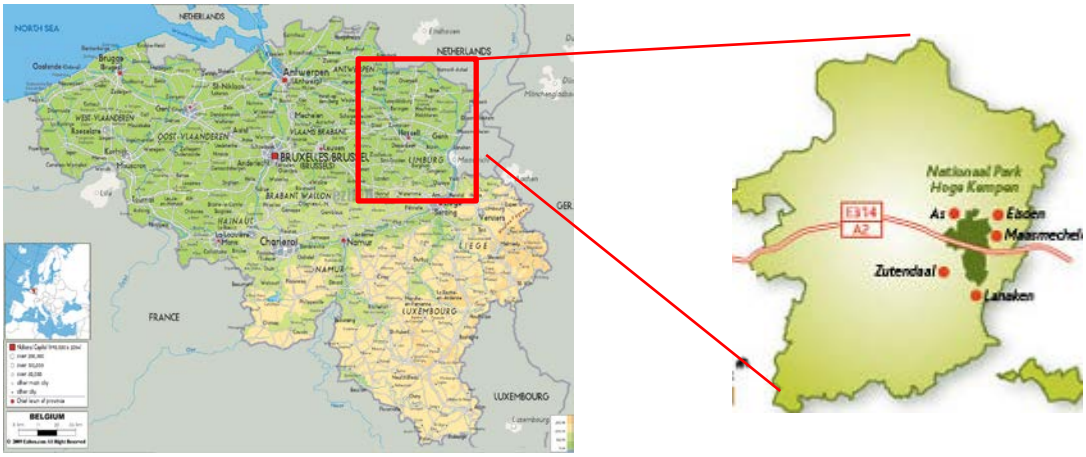


Figure 1. Situation of the NPHK (source: elizon maps)

Like the majority of protected areas worldwide, the NPHK relies largely on governmental budget for ES and biodiversity conservation, habitat restoration and visitor management. In the European Union, financing the costs to achieve these objectives is a highly debated political issue. Further empirical information is necessary to demonstrate public preferences for different management options of protected areas (Hoyos *et al.*, 2012). Although hypothetical, a stated preference survey is considered to elicit the socio-economic benefits - or Total Economic Value (TEV) - of the park, and assist in the further development of conservation payment mechanisms. The economic valuation implies that stakeholders need to be consulted to express their preferences for the park's characteristics that may be included in the CE. Park managers, tourism businesses, residents, regional and governmental agencies, private land owners, and nature organisations are among the relevant park stakeholders. Stakeholder participation for assessing the needs of the local community has been frequently recommended and named a success factor of natural resource management and protected area tourism (Millennium Ecosystem Assessment, 2005, Dudley, 2008, Liu *et al.*, 2010, Maynard *et al.*, 2014). Combining the stakeholders' wide-ranging objectives and their varying level of awareness and knowledge about ES and biodiversity is a key milestone in successful park management (Dudley, 2008, Darvill and Lindo, 2015). Consequently, the diversity of stakeholder perceptions, knowledge, and representation of management options (*e.g* increase vs maintenance of current visitor facilities, conservation interventions, accessibility) needs to be reflected in the CE.

3. FRAMEWORK FOR THE IDENTIFICATION AND SELECTION OF CE ATTRIBUTES AND ATTRIBUTE LEVELS

There exists no 'gold standard' for the generation and definition of attributes (Louviere *et al.*, 2000). It is generally acknowledged that, in order to be policy relevant, attributes for a CE should be chosen in such a way that they reflect the perspectives of both the public (affected by some quantitative or qualitative environmental change) and the resource managers. Moreover, CE attributes ideally have a sound scientific basis and can be translated from scientific indicators or parameter values to understandable and meaningful descriptions of the relevant components at hand (Lancaster, 1966, Blamey *et al.*, 2002, Ryan *et al.*, 2009, Abiiro *et al.*, 2014). Researchers should clearly consider two phases of attribute development: 1) initial attribute identification through qualitative research, 2) the generation of meaningful wording and the final attribute selection, which inevitably incorporates the identification of attribute-levels (Coast *et al.*, 2012). Phase one - the identification of initial attributes - should be informed by a thorough literature review on the study topic to obtain attributes which can potentially be included in a CE. Identifying attributes exclusively on the basis of a literature review may lead to the omission of attributes considered important and relevant by the target population (Coast *et al.*, 2012). Consequently, a rigorous qualitative study is required to identify initial attributes important to stakeholders and contributes to the reduction and misspecification of attributes and invalid CE estimates (Coast *et al.*, 2012, Abiiro *et al.*, 2014). Phase two, the final attribute selection, requires the support of experts in order to identify mutually exclusive and measurable attributes and attribute-levels.

Several studies in health economics (Coast and Horrocks, 2007, Coast *et al.*, 2012, Kløjgaard *et al.*, 2012, Abiiro *et al.*, 2014, Michaels-Igbokwe *et al.*, 2014) and one in ecological economics (Armatas *et al.*, 2014) highlight the need for more rigorous and thorough reporting of the set of relevant attributes and levels to be accounted for in stated preference surveys and the current lack thereof. They argue that the most frequently used qualitative methods, such as FGs and INTs, are usually poorly described when employed for attribute selection in CE in stated preference research. We respond to the call expressed in these studies by presenting a methodological framework for the attribute and level selection process (Fig. 2). The semi-qualitative protocol proposed in this paper is divided into five distinct stages, which are further detailed in the 'Methodology' and 'Results' section.

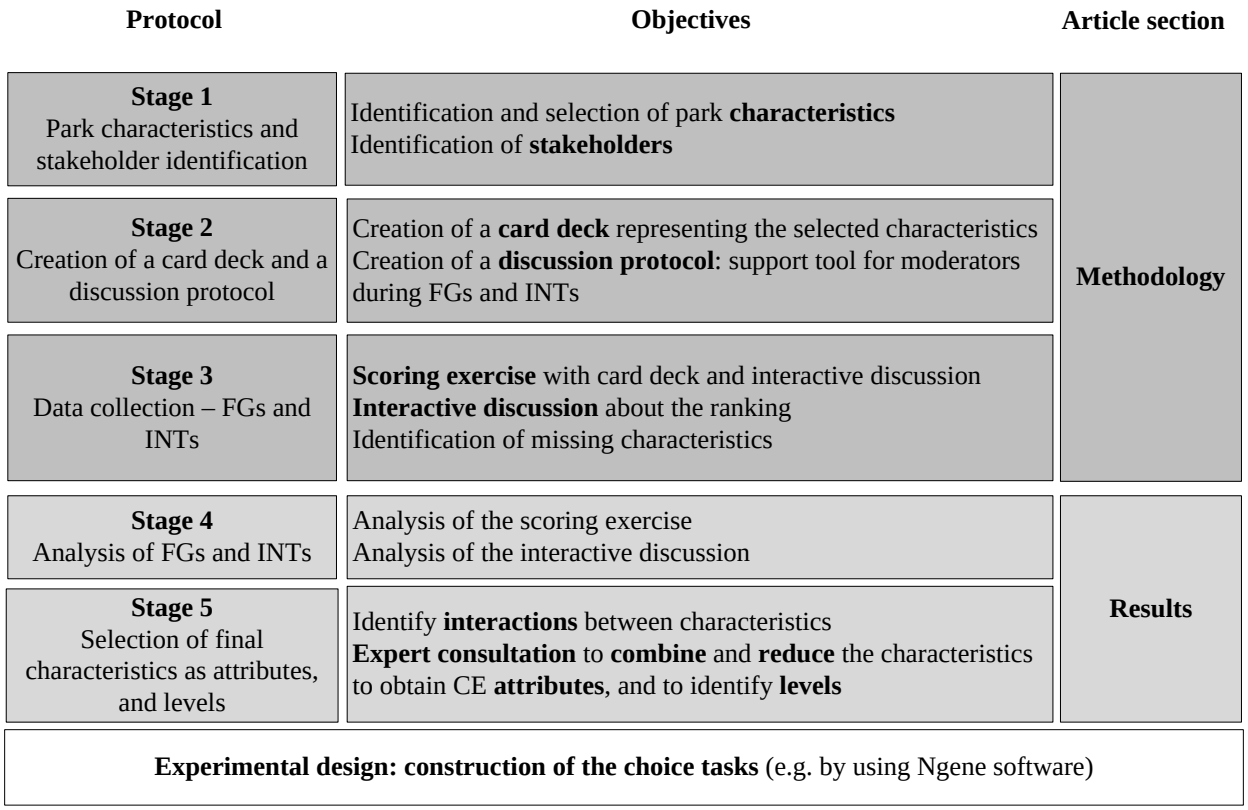


Figure 2. Methodological framework of the attribute and level selection process (Based on Coast and Horrocks, 2007; Klojgaard et al., 2012; Coast et al., 2012; Michaels-Igbokwe et al., 2014; Abiiro et al., 2014 and Armatas et al., 2014).

The use and usefulness of the methodological framework is operationalized and illustrated in our case study. In stage 1, we identify park characteristics from a literature review and consultations with researchers and park officers. Moreover, park stakeholders are identified. Stage 1 serves as a basis for stage 2 where a picture-based card deck and a practical discussion protocol are designed as support for the FGs and INTs. During the semi-qualitative process in stage 3, the pre-selected characteristics are scored, ranked and discussed by the stakeholders. In stage 4, the rankings, tapes and notes are analysed, while in stage 5 the final attribute selection takes place. Here, we label characteristics as attributes once they can be inserted in the CE. For the purpose of this study, we focus exclusively on the attribute and level selection process. One can refer to Reed Johnson, F., *et al*. (2013) for more detailed information about experimental design options.

4. METHODOLOGY: ATTRIBUTE SELECTION PROCESS

Stage 1: Identification of park characteristics and stakeholders

A preliminary list of park characteristics was developed using official ES classifications (MEA, CICES and de Groot *et al.*, 2002). Grey literature, books and reports on the study area were examined. The selection process necessitates rigour and extended expert consultation to ensure that the most relevant and representative characteristics of the park are included. Three pilot group discussions consisting of respondents who were unfamiliar with the park (scholars, external relations, potential park visitors), three expert meetings with park managers, and four meetings with university researchers (ecologists and economists) were held to improve definitions of park characteristics and to identify additional ones that were not encountered during the literature review. Based on these consultations, we identified a list of 50 characteristics: 22 cultural services, 6 regulating services, 3 provisioning services, 3 supporting services, 6 land use types, 10 infrastructure-related characteristics (Tab. 1). This list was not restricted to ES *per se* in order to avoid miscomprehension of the respondents given that a majority is unfamiliar with the ES concept (Barkmann *et al.*, 2008, Armatas *et al.*, 2014). Also, a protected area possesses non-ES characteristics that stakeholders may care about, such as accessibility and tourist attractions.

Table 1. List of 50 park characteristics

Category	**Based on literature review, expert meetings with park managers and university researchers**
	Park characteristics
Cultural services	
Recreational activities	Biking, Hunting, Fishing, Wildlife spotting, Mountain biking, Walking, Jogging, Dog walking, Horseback riding, Hunting (management), Presence of Wild boar
Scenery	Landscape variety, Panoramic view
Heritage	Cultural heritage
Research & Education	Research opportunities, Environmental education, Youth activities
Spiritual	Peacefulness, Nature conservation, Social interactions
Regulating	Groundwater storage, Carbon storage, Pollination, Water purification, Air purification, Control of invasive species
Provisioning	Wood, Sand, Drinking water
Habitat/ supporting	Biodiversity refuge, Soil fertility, Park size
Land use type	Heathlands, Natural forests, Wetlands, Pine forests, Agricultural land, Restored sites
Infrastructure/services	
Tourism	B&Bs, Hotels, Camping & Holiday park, Local products, Gateway attractions, Art, Pick-nick opportunities
Park	Connectivity, Accessibility, Signposting, Wildlife crossings, Grazing horses

For the FGs and INTs, stakeholder participants were selected by maximum variation sampling: participants were selected using internet-based research of the study area, meetings with park managers and snowball sampling during the actual data collection. Like Armatas et al. (2014), we aimed at a diversity rather than a quantity of opinions by surveying 46

stakeholders. 46 appeared to be a sufficient number as saturation was reached, *i.e.* the top characteristics appeared to be essentially the same across stakeholder groups.

Stage 2: Creation of a card deck and a discussion protocol

Stage 2 consisted in presenting the characteristics in a way that is easily understood by respondents who are unfamiliar with ES and biodiversity features. The often limited knowledge of people of ES forced us to seek out a way of translating the ES concept using appropriate local images to fit the local context. Instead of using statements like in the Q-methodology, we favoured a picture-based approach in order to facilitate the interpretation of each service. Validity of using photographic representations to assess perception of nature was established by different studies (Martín-López *et al.*, 2007, Gómez-Baggethun and Kelemen, 2008, Kenter *et al.*, 2011, Abiiro *et al.*, 2014).

Building upon approaches presented in Armatas *et al.* (2014) and the information provided by the Belgian Institute for Nature and Forest (INBO, 2014), a 50-picture-card deck was created to obtain a representation of relevant characteristics from the viewpoint of the park stakeholders (Fig. 3). People seem to grasp the concepts faster and more clearly when visual examples are given, being on screen, on pictures or on maps (Petheram and Campbell, 2010). It also reduces the cognitive demand on respondents (Cerda *et al.*, 2013). Most cards presented photographs from the NPHK and were randomly numbered from 1 to 50.

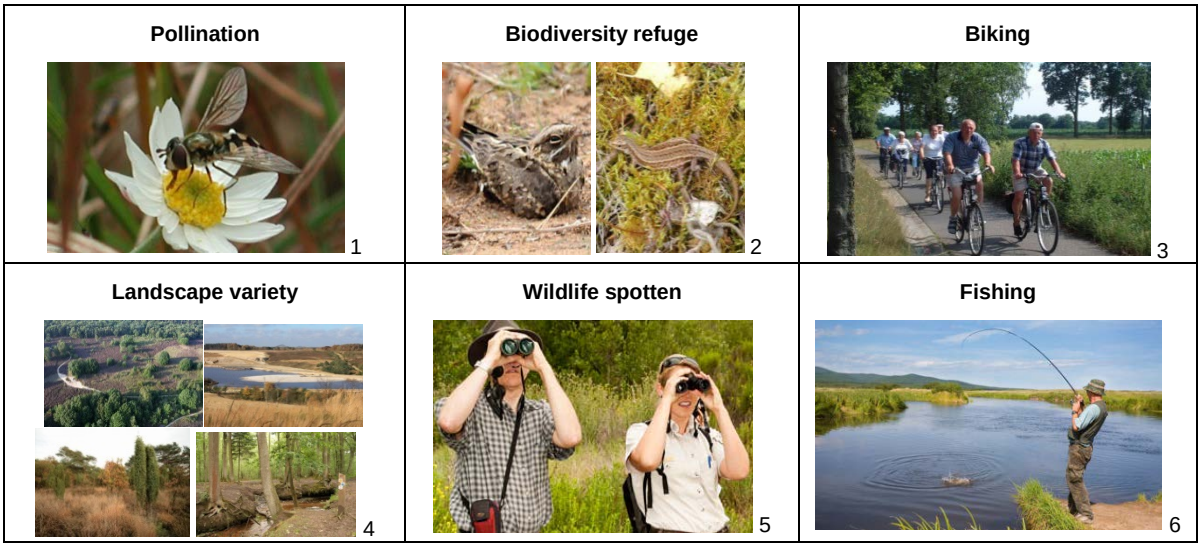


Figure 3. Examples of cards used in the scoring exercise

The card deck serves as a visual support for a scoring exercise to be completed during the FGs and INTs (Stage 3). An 4-page protocol was designed for the group facilitators and interviewers. It includes a description of the roles (moderator, assistant), the instructions of the exercise and the discussion, and of the final attribute selection (see Appendix A). This protocol ensured that moderators become familiar with all the necessary steps. Having a

detailed protocol is particularly important if the moderation is the task of someone other than the leading researcher and guarantees uniformity across all FGs and INTs (Krueger and Casey, 2000).

Stage 3: Data collection - focus groups and interviews

Between October 2014 and January 2015, a semi-qualitative approach involving four FGs and thirteen INTs were conducted. INTs were conducted with a single participant either in Dutch or English. In total, 46 stakeholders participated in the FGs and INTs (Tab. 2). Stakeholders were grouped into 9 different background categories and consisted of: 25 males and 21 women; 31 respondents were aged 46 and above; 26 participants live in one of the municipalities surrounding the park (park residents) and 20 in the rest of the province of Limburg in which the NPHK is located. 18 local residents have lived and worked in the park's vicinity for over 20 years (Tab. 2).

For the FGs, the active participation process entailed two main parts: 1) a scoring exercise and 2) an interactive discussion. The first part aimed at scoring the 50 characteristics on a scale from 'very important' (score 1) to 'not important at all' (score 5), and hence, at obtaining a list of the most relevant characteristics for the respondents. For FGs, the moderator was assisted by a note taker and a computer manager. A slide show consisted of three introductory slides, with information on the purpose of the overall study and the need for stakeholder consultation, and the card deck. The projected cards were linked to the polling software TurningPoint. This software is a powerful way to collect data quickly from large groups of participants. It actively involves participants and keeps them engaged and interested in the presentation material. Each participant had a voting remote to assign a score to each card so as to state how important the characteristic represented on the card is to their own well-being; hence providing their individual preferences. Since participants were expected to provide their individual preferences, we ensured that there was no or very limited interaction during the scoring exercise throughout the FGs. The meaning of the cards was briefly described to have uniform understanding and prevent participants from having varying interpretations. This exercise resulted in a ranking of the characteristics, based on the mean scores of all group participants. In the INTs, the respondent had the 50-picture-card deck in one hand and placed the cards one by one onto one of the five categories of importance.

Table 2. Profile of the participating stakeholders (N=46)

Background category	N	Focus group (FG) Interview (INT)	Gender	N	Age	N	Residence	N
Park office managers	6+1	FG + INT	Males	25	18-45	15	Park municipalities	26
Tourism enterprises	6+2	FG + INT	Females	21	46-55	20	Since 0-10 years	2
Local government	10	FG			56-65	4	Since 10-20 years	6
Regional government	1	INT			65+	7	Over 20 years or always	18
University/Research	2	INT					Other municipalities in the province	20
Coordinator of the private land owners	1	INT						
Industry	2	INT						
Nature organisation	11+2	FG + INT						
Bikers	2	INT						

In the second part, FG participants discussed the ranking and were given the opportunity to make changes and to identify potentially important missing characteristics. This interactive discussion further informed us about the respondents' underlying reasons and motivations for the selection of the park characteristics and about their level of understanding. Like the discourse-based valuation approach suggested by Wilson and Howarth (2002) we encouraged participants to agree upon the group ranking in order to elicit consensus-based preferences for the most important park characteristics. A consensual decision was desired, but unanimous consensus was not deemed to be an imperative. An ideal group for participatory democracy is a group where participants have shared values, behave like friends and work at preserving equality (Catt, 2002, Lo, 2013). We assumed a priori that a broad consensus would be reached due to the homogeneity in background of the groups. The group homogeneity was the reason for conducting a prioritisation process that includes a group review of the aggregate score of the individual rankings (Hiligsmann *et al.*, 2013). The group reached consensus upon the most important characteristics for the well-being of the park community, by agreeing on the fifteen most important characteristics (Eagles *et al.*, 2013). The limited number of fifteen characteristics was chosen due to the homogeneity of answers, the choice set design constraints and to facilitate the final attribute selection. CE studies recommend the inclusion of between 4 and 8 attributes (Coast and Horrocks, 2007). The number of top-ranked characteristics displayed can vary. However, we suggest not to show beyond fifteen to twenty characteristics to avoid fatigue and disinterest of the participants. In INTs, only the characteristics categorised as 'most important' (and 'important' in case very few cards were placed in category 'most important') were kept to be further discussed.

After agreeing on the ranking, FG and INT respondents were asked to specify which from the pre-selected most important characteristics (15 in the FGs) required intervention by responding '*Yes*' or '*No*' to the question "S*hould there be more of this characteristic?"*. Answers inform us where improvements should occur, and thus what changes to potentially include in the CE. Stakeholders can find a characteristic very important and be satisfied with its current state, meaning that no change or action is required. If they express no willingness for change, a characteristic should not be included in a CE. '*Not applicable*' was a possible answer when the situation of a characteristic was unknown or not quantifiable, such as 'groundwater storage'.

After the participation process, the answer '*Yes*' was coded as 2 points and '*No*' or '*Not applicable*' as zero points. When no agreement was found, the answer '50/50' was coded as 1 point. In the end, if the sum of the points of all groups was higher than the half, it meant that the majority of stakeholders would like to see more of this characteristic. For instance, in our case the maximum sum of points was 18 (2 points x 9 stakeholder groups). So all top characteristics with a sum higher than 9 were retained as CE candidates. Because the 50/50 option could not be produced by INTs, we recoded the 50/50 as zero to test if the results are still robust; it was the case. A step-by-step protocol is available in Appendix A.

FGs and INTs took on average 1h30 (30 minutes polling, 1 hour discussion) and 45 minutes respectively. Each participant filled out a consent form that included background socio-demographic questions (age, occupation, residence, work place, visit frequency to the park, and residence duration if residing in one of the six municipalities surrounding the park). FGs were analysed using note-based and INTs using tape-based analysis. INTs were partially transcribed. The notes from the FG discussions were used as abridged transcripts of the most relevant and useful comments and viewpoints raised during the interactive discussion. Transcribing consisted in working out the INTs from the recordings and writing them out. We opted for abridged transcripts due to the researchers' thorough understanding of the purpose of the study (Krueger and Casey, 2000). Moreover, only the comments related to the cards presented were retained. An abridged transcript is much shorter (*i.e*. it focuses exclusively on important portions) than a full transcript in a transcript-based analysis (entire discussion).

5. RESULTS

Stage 4: Analysis of the focus groups and interviews

Analysis of the scoring exercise

The ranking lists of all FGs and INTs were obtained by simple arithmetic mean (Kaplowitz and Hoehn, 2001). The final ranking is based on the aggregated mean scores of all stakeholder groups with each stakeholder group having the same weight (Tab. 3). Differences in perceptions regarding the importance of park characteristics were evaluated across stakeholder categories. Differences were tested with a non-parametric Kruskal-Wallis test using IBM SPSS Statistics 22 (Cárcamo *et al.*, 2014). Given the small sample sizes, we use p-value < 0.1 as a cut-off point.

Considering the mean score of all stakeholders, 6 characteristics achieved a score of 'very important' (mean score 1-1.5), 23 of 'important' (1.51-2.5), 16 of 'neutral' (2.51-3.5) and 3 were 'less important' (3.51-4.5). For the 'very important' characteristics, differences among stakeholder organisations are found for 'nature conservation' ($p < .06$), 'natural forests' ($p < .059$), 'biodiversity refuge' ($p < .013$), heathlands ($p < .015$) and landscape variety' ($p < .08$). When analysing pairwise comparisons, we observe that the industry (IND) and the bikers (BIK) were the groups that most affected the p–values, thus by attributing lowest scores. We are aware that this result can be due to the small size of these groups (2 individuals). There were no differences in perceptions between the respondents living around the park and those further away, except for the characteristic 'signposting' (Mann-Whitney test, $p < 0.016$).

Translating the scores of the characteristics to the importance of the categories, the 'very important' scores were assigned to 'land use types' (natural forests = 1.17, wetlands = 1.37, heathlands = 1.46), 'cultural services' (nature conservation = 1.17, landscape variety = 1.43), 'supporting services' (biodiversity refuge = 1.2) and 'regulating services' (air purification = 1.5). Characteristics ranked 'very important' by all stakeholder organisations can be judged relevant for inclusion in a valuation survey and are likely to be valued by respondents. Characteristics perceived as rather neutral should not be candidates for inclusion in a CE (Armatas *et al.*, 2014), since they are likely to be of little concern to respondents (*i.e.* tourism accommodations, art and wood). 'Not important at all' characteristics may also be candidate for a CE, in particular for researchers wanting to identify protest bids (Greiner *et al.*, 2014). Fishing, sand and hunting (as a hobby) obtained the most negative scores, being considered "inadequate in a national park".

Table 3. Stakeholders' perceptions regarding the importance of the NPHK's characteristics.

Characteristic	Mean		Kruskal-Wallis test		H_0=distribution is the same across groups	Characteristic	Mean		Kruskal-Wallis test		H_0=distribution is the same across groups
	Score (1-5)	SD	χ^2 (df=8)	P-value			Score (1-5)	SD	χ^2 (df=8)	P-value	
Nature conservation	1.17	0.437	14.936	.06	**No**	Panoramic view	2.26	0.929	12.796	.119	Yes
Natural forests	1.17	0.437	15.028	.059	**No**	Soil fertility	2.35	0.059	5.985	.649	Yes
Biodiversity refuge	1.2	0.401	19.314	.013	**No**	Wildlife spotting	2.39	0.881	6.665	.573	Yes
Wetlands	1.37	0.572	13.328	.101	Yes	Pic-nick	2.39	0.802	11.591	.170	Yes
Landscape variety	1.43	0.688	14.053	.08	**No**	Youth activities	2.48	1.027	9.742	.284	Yes
Heathlands	1.46	0.959	18.992	.015	**No**	UNESCO	2.5	1.394	17.877	.022	**No**
Air purification	1.5	0.624	10.846	.211	Yes	Grazing horses	2.52	1.15	13.153	.107	Yes
Education	1.52	0.809	8.052	.428	Yes	Local products	2.52	1.027	10.281	.246	Yes
Walking	1.54	0.836	13.334	.101	Yes	Social interactions	2.53	0.894	12.139	.145	Yes
Peacefulness	1.59	0.858	11.57	.171	Yes	B&Bs	2.63	0.951	24.778	.002	**No**
Water purification	1.72	0.861	11.574	.171	Yes	Pine forests	2.78	1.134	2.937	.938	Yes
Wildlife crossings	1.74	0.976	11.692	.165	Yes	Hotels	3	1.135	16.224	.039	No
Connectivity	1.76	0.822	2.393	.967	Yes	Wild boar	3.02	1.273	9.021	.341	Yes
Restored sites	1.76	0.993	8.293	.405	Yes	Hunting (management)	3.09	1.226	7.615	.472	Yes
Groundwater storage	1.8	0.885	13.126	.108	Yes	Jogging	3.15	1.074	20.335	.009	**No**
Carbon storage	1.83	0.926	5.799	.670	Yes	Horseback riding	3.3	1.072	17.590	.025	**No**
Pollination	1.85	0.894	15.985	.043	**No**	Camping	3.3	1.364	22.700	.004	**No**
Park size	1.96	0.893	12.615	.126	Yes	Art	3.39	1.000	12.510	.130	Yes
Source drinking water	1.96	1.074	12.713	.122	Yes	Mountain biking	3.41	1.147	9.999	.265	Yes
Signposting	1.98	0.954	19.533	.012	**No**	Agricultural land	3.41	1.275	8.786	.361	Yes
Accessibility	2.04	1.032	18.809	.016	**No**	Wood	3.43	1.128	13.421	.098	**No**
Biking	2.04	1.032	20.13	.01	**No**	Dog walking	3.43	1.259	10.275	.246	Yes
Cultural heritage	2.11	0.875	9.581	.296	Yes	Fishing	3.8	0.957	4.610	.798	Yes
Gateway attractions	2.13	1.204	13.999	.082	**No**	Sand	4.04	1.095	12.719	.122	Yes
Research opportunities	2.17	0.950	8.233	.411	Yes	Hunting (hobby)	4.3	1.072	7.300	.505	Yes

Values are the mean of all stakeholders + standard deviation (SD). A Kruskal-Wallis test was used to calculate the differences among stakeholder groups.

Analysis of the interactive discussion

As hypothesised, the homogenous FGs did not raise strong contradictory arguments. Even in the case of diverging views, the final ranking did not change. Like in every group exercise, some participants expressed themselves more often than others, but the moderator ensured that everyone one had a word to say. All four FGs reached consensus upon the fifteen most important characteristics and agreed that it reflected their individual scores. Consensus was easily reached and agreed upon, most likely because FGs were homogenous. Moreover, each group meets on a regular basis, so participants knew each other relatively well.

Regarding the question "*Should there be more of this characteristic?"*, the addition of the points of all groups shows that the majority of respondents answered '*Yes*' to six characteristics (Tab. 4). The maximum of points that could be allocated to a characteristic was 18 (2 points x 9 stakeholder groups). The six selected characteristics were 'natural forests' (16 points), 'biodiversity refuge' (13 points), 'nature conservation' (11 points), 'heathlands' (11 points), 'air purification' (10 points) and 'education' (10 points). Bikers exclusively expressed a willingness for improvement regarding biking trails and opportunities; therefore only zeros are presented in Table 4.

Table 4. Answers to '*Should there be more of this characteristic?*'; 2 = Yes, 1 = 50/50, 0 = No or Not applicable; the black cells are for the characteristics that were absent from the top 15.

	Park office managers	Tourism enterprises	Local government	Regional government	University/ Research	Coordinator of private land owners	Industry	Nature organisation	Bikers	**Total**
Nature conservation	2	1	2	0	2	2		2	0	**11**
Natural forests	2	2	2	2	2	2	2	2	0	**16**
Biodiversity refuge	2	1	2	2	2	2	0	2	0	**13**
Heathlands	2	1	0	2	2	2		2	0	**11**
Air purification		2	2	0	2	2	0	2	0	**10**
Education	0	2	0	2	2		2	2	0	**10**

The transcripts and notes revealed the rationale for the respondents' choices and that certain characteristics are subject to different interpretations, varying among stakeholder groups. For instance, stakeholders from the industry and the sport sector assign less importance to nature conservation and aesthetics. Although their opinion may not be aligned with the objectives of the protected area, their opinions as stakeholder of the community need to be integrated. The transcripts also reveal policy-relevant attributes that are not necessarily demand-relevant, which is beneficial considering the need to include attributes that are both demand- and policy-relevant. For instance, 'pine forests' is not demand-relevant for most stakeholder organisations, but is of policy and management concern for the NPHK, since these forests are

commercially exploited. For the top-ranked characteristics, here are some arguments respondents used to support their choice:

> Environmental education is very important, in particular among the young generation. It is a main objective and a basic component of a national park. (Regional government)
>
> If it is absent there is no continuity. (Local government)
>
> The presence of grazing horses and the control of invasive species are necessary management practices that contribute to some of the 'very important' park characteristics such as heathland, landscape variety, and biodiversity refuge. (Coordinator of private land owners)
>
> The objectives of the NPHK are recreation and conservation. (Biker)

Linkages between characteristics were pointed out. Such observations are particularly relevant knowing that CE attributes need to be mutually exclusive:

> The top-4 characteristics are very linked to each other: heathlands, biodiversity, landscape variety, natural forests. (Park office managers)

Regulating ES appeared more difficult to understand. Respondents often showed a lack of knowledge regarding their role and function in the park:

> It is difficult to assess and quantify the impact of trees on air purification. This service is more important in urban areas. The paradox is that to improve this ecosystem function, pine trees should be safeguarded. (Research)

These discussion outcomes show that the wording and the presented information require very careful selection in case regulating ES are to be presented. Although respondents were asked if any important characteristics were missing, no new elements emerged. Some respondents mentioned certain characteristics as being important policy items, such as the creation of more unmanaged broad-leaf forests, wildlife disturbance, monitoring of groundwater level, and the control of invasive species. Furthermore, stakeholders perceived interrelations between several characteristics and perceived the three categories of ES in different ways. For example, they were unanimous regarding the important role of the NPHK's cultural services, like nature conservation and landscape variety, thus explaining the very frequent appearance of these characteristics together in top of the ranking list which is similar to the findings of Martin-Lopez *et al.* (2012). Finally, we would like to note that, except provisioning services, all categories of ES were represented.

Stage 5: Selection of final characteristics as CE attributes, and attribute-levels

The attribute selection depends on the study topic and the results the researcher aims to obtain from the CE. In our case, we will calculate the willingness-to-pay of the park visitors for the ES and biodiversity of the NPHK**.** This stage consists in combining and reducing the top-ranked characteristics to a manageable number of attributes (4-8) within a CE (Abiiro *et al.*, 2014). This final attribute and attribute-level selection is based on the outcomes of the scoring

exercise and the interactive discussion, the transcripts, inclusion-exclusion criteria, previous CE studies, and an iterative expert and stakeholder consultation process. To deal with the challenging attribute combination, reduction and level selection tasks while meeting the inclusion-exclusion criteria, we opted for an iterative process with experts (Fig. 4).

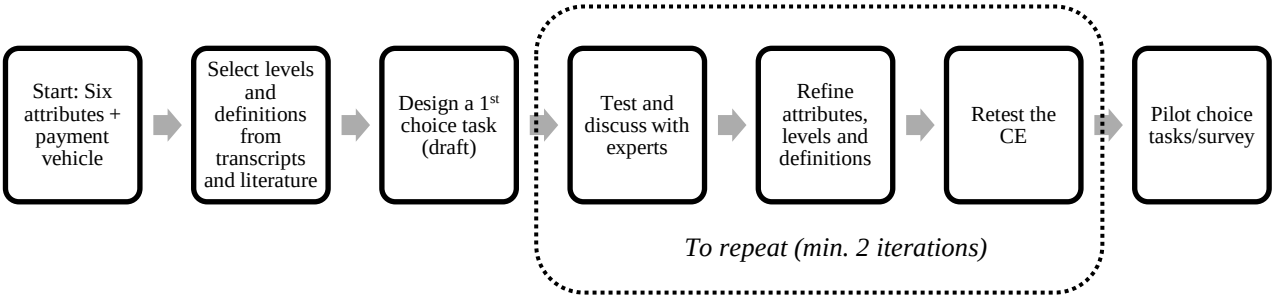


Figure 4. Iterative process leading to the final selection of attributes and attribute-levels

Furthermore, the selection process is guided by the following inclusion-exclusion criteria, whereby attributes for environmental CE should be:

- not overlapping other attributes (to avoid inter-attribute correlation) / mutually exclusive in nature (Coast *et al.*, 2012, Abiiro *et al.*, 2014)
- policy-relevant (Blamey *et al.*, 2002, Ferrini and Scarpa, 2007)
- demand-relevant: important, understandable and meaningful to people and relate to their reasons for having willingness-to-pay to conserve biodiversity and ES (Blamey *et al.*, 2002, Lancsar and Louviere, 2006)
- measurable (Blamey *et al.*, 2002, Zander and Drucker, 2008)
- ecologically and economically relevant (Johnston *et al.*, 2012)
- describable by combining simple explanations and visual instruments such as photographs, charts, and pictures (Cerda and Losada, 2013)
- limited to a number between 4 and 8 because trade-offs become difficult to understand (Abiiro *et al.*, 2014)

A total of 14 people (park managers, biologists and economists) guided the choice of attributes, levels, and the payment vehicle. We opted for an entrance fee because it represents an important source of conservation funding for protected areas worldwide. Many studies have revealed that visitors were willing to pay more than the current fees for biodiversity conservation and environmental protection (Wang and Jia, 2012, Kaffashi *et al.*, 2015). In the case of the NPHK, an entrance fee is already asked at one gateway, and thus is politically and socially accepted. The 14 people were first presented with a draft choice task containing attributes representing the six characteristics: nature conservation, natural forests, biodiversity

refuge, heathlands, air purification, environmental education. An entrance fee as payment vehicle was added to the choice task (Tab. 5). Although it was ranked 4th in the scoring exercise, the characteristic 'wetlands' was excluded because "*They represent small scattered surfaces in the heathland landscape, and so these two land use types can be presented together*" *(Park office manager)*. The second reason for the exclusion of 'wetlands' and 'landscape variety' is that stakeholders were satisfied with their current status. The characteristics obtained respectively a total of 9 and 7 points (Tab. 4), thus not reaching a majority of votes in favour of 'more of this characteristic. 'Natural forests' and 'heathlands' overlap with 'landscape variety'; therefore the latter is not visible in Table 5.

The selection of initial levels was based on previous CE studies, grey literature and reports from the NPHK, and the inclusion-exclusion criteria. The 'air purification' attribute proved more difficult to assign a level at first. A number of CE have assessed willingness-to-pay values for air quality or the mitigation of air pollution, typically in highly polluted areas and included health-related attributes. Most of our respondents care about air quality, but did not raise any health issues. Knowing that units and indicators of air quality are not easily understood (*e.g.* Particulate Matter PM 2,5, PM 10), we kept the identification of 'air purification' levels open for discussion with the experts.

Throughout this iterative process, the choice task was modified and improved five times in order to obtain the pilot choice tasks. Although experts agreed unanimously upon the chosen attributes, it became clear that seven attributes were too many. In Table 5, arguments for inclusion or modification in relation to the first choice task are provided. Attributes that overlapped with other attributes were removed or merged in order to avoid inter-attribute correlation, such as the land use types with 'biodiversity refuge'. This observation is consistent with the park managers' statement "*The top-4 characteristics are very linked to each other: heathlands, biodiversity, landscape variety, natural forests*."

Table 5. Attributes and initial attribute levels presented to experts iteratively for the final inclusion

Characteristic	Attribute	Levels	Argument	Reasons for inclusion in the CE	
				Final ranking position	**There should be more** (max. points: 18)
Nature conservation	Chances of observing the red deer	0 1 out of 50 walks 2 out of 50 walks	"What is missing in the NPHK are large mammals, which greatly foster people's interest and are used as flagship species"	Ranked 1st *Mean score: 1.17*	11
Natural forests	Natural forests	5%, 10%, 15% of total park surface	"They provide clean air, peacefulness and at global scale are an important carbon sink."	Ranked 1st *Mean score: 1.17*	16
Biodiversity refuge	Biodiversity	0%, 10%, 20% increase	"Biodiversity needs to be included, but respondents do not know what a percentage increase means."	Ranked 2nd *Mean score: 1.2* Then merged with 'natural forests' and 'heathlands'	13
Heathlands	Heathlands	20%, 30%, 40% of total park surface	"The percentages are realistic, but your park coverage is not complete when considering the 'natural forest' attribute'. Also, 'biodiversity' is correlated with land use types, so you may want to combine all these items."	Ranked 6th *Mean score: 1.46* Then merged with 'biodiversity' and 'natural forests'	11
Air purification	Air purification or clean air	Open for discussion	"It is difficult to assess and quantify the impact of trees on air purification. Enlarging the park surface area would improve the provision of this ES. "	Ranked 7th *Mean score: 1.5*	10
Education	Environmental education a) Availability of rangers b) Fixed activities	a) 3 weeks, 1 week or a few days in advance b) 2 Sundays/month, 4 Sundays/month or every weekend day	"This is very clear and straightforward. You may want to add the opening hours of the gateways. They would like to be open all year long, but financially it is currently not feasible."	Ranked 8th *Mean score: 1.52*	10
	Entrance fee	0, 2, 4, 6, 8 €	"From an experimental design point of view, I would add more levels."		

This final selection process led to the retention of 4 attributes with 3 levels and a payment vehicle attribute with 6 levels (Tab. 6). In the ultimate iteration, all participants agreed that the selected attributes and levels satisfy the CE inclusion criteria, *i.e*. they reflect the characteristics of the NPHK, are perceived as important by the stakeholder community, are understandable, and mutually exclusive. Health economics papers were helpful with guiding the final attribute and attribute-level stage, since they clearly described how they made use of the transcripts and the inclusion criteria (Coast and Horrocks, 2007, Hiligsmann *et al.*, 2013, Abiiro *et al.*, 2014).

Table 6. Final selection of CE attributes, levels and definitions

Attribute	Description	Levels	Description
Biodiversity	An increase in the surface of heathland, wetland and oak-birch forest would provide more space and shelter for rare and endangered species. The surface in pine forests would decrease in favour of the other land use types.	*pine forest: 60%* *natural forest: 10%* *heathland: 30%*	*Current situation*
		pine forest: 45% natural forest: 20% heathland: 35%	Medium benefits: Less timber More peacefulness More space for heathland and natural forest biodiversity Less invasive species, A more balanced landscape
		pine forest: 30% natural forest: 30% heathland: 40%	High benefits: The above benefits occur at a higher degree
Chance to observe the Red deer	The NPHK is an appropriate environment for the return of the Red deer. Reintroduction is an eventuality.	*No chance* 1 out of 10 walks 2 out of 10 walks	*Current situation* 10% chance 20% chance
Environmental education	Educational activities aim to increase environmental awareness and knowledge among all generations : a. Ranger availability b. Fixed activities c. Opening hours of the gateways	*a. 3 weeks in advance* *b. 2 Sundays/month* *c. closed 1 day/week*	*Current situation*
		a. 1 week in advance b. 4 Sundays/month c. Open every between April and November	Medium benefits: The feeling of learning, awareness of the importance of nature, valuable exchange with other visitors and rangers
		a. A few days in advance b. Every weekend day c. open all year long	High benefits: The above visitor services occur at a higher degree.
Air purification	Woodlands and vegetation in general improve air quality. Trees and plants capture harmful fine dust from the air.	*6.000 ha* 8.000 ha 10.000 ha	*Current situation* 1/3 larger 2/3 larger
Entrance fee	Price for a day visit in the NPHK. The revenues flow directly to the park conservation fund.	*0 €* 2, 4, 6, 8, 10, 12 €	*Current situation* Hypothetical entrance fees

6. DISCUSSION

This paper contributes to the literature on attribute and level selection for CE by designing a transparent and user-friendly semi-qualitative protocol. Moreover, it does so for an environmental context, which often deals with unfamiliar public goods. The protocol we propose comprises a picture-based scoring exercise (Likert-scale type) followed by an (interactive) discussion with park stakeholders. It results in a ranking of most to least important park characteristics according to stakeholders, as well as in abridged transcripts that revealed the rationale for the ranking. Consequently, this input is used during iterative discussions with experts and stakeholders and results in the final selection of CE attributes, levels and sound, yet understandable definitions. Hence, compared to previous literature on environmental CE our approach covers the complete selection process prior to the choice task design.

The entire selection process presented in this paper can be qualified as semi-qualitative , that is one with concise structured outputs rather than wordy transcripts. This process has the advantage that the output is apparently simple and the process short (Forrester *et al.*, 2015). Our protocol can be confronted with several methods. The large majority of studies reporting a detailed and transparent attribute and attribute-level selection process are found in health economics. FG and INT participants generally belong to one or two specific groups (*e.g.* osteoporosis patients, young people aged 15-24, health workers), which makes the sampling and gathering of stakeholders easier than in an environmental context where people with very diverse backgrounds need to be consulted. Unlike the often intangible nature of environmental attributes, health CE include attributes that are observable, easy to understand, and directly affecting the respondents (*e.g.* specific medicines, treatments, insurance regulations). In ecological economics, we are not aware of any publication proposing a detailed and easy-to-use selection process that includes the identification of both attributes and attribute-levels. One study applies the nominal group technique to prioritise attributes for inclusion in CEs for drug treatment for osteoporosis (Hiligsmann *et al.*, 2013). This technique is composed of a ranking exercise (ranking 12 previously selected attributes) and a follow-up discussion. Similar to our findings, the nominal group technique discussions did not substantially affect rank order of preferences for the attributes in the total group when compared with rank order before the discussion, indicating agreement for the most important attributes. Unlike this method, our protocol allows the treatment of a large amount of characteristics to obtain a selection.

Another popular elicitation method is the Q-methodology (Q). Our approach resembles the Q in the way that it is particularly suited to purposive sampling of individually held perspectives within stakeholder groups (Forrester *et al.*, 2015). The Q involves the sorting of items and the

comparison of opinions between stakeholder groups. Researchers have applied the Q for learning how to design policy solutions and to improve decision-making in instances of conflict-ridden ecosystem management issues (Swedeen, 2006, Forrester *et al.*, 2015). Similarly, it provides a transferable, structured and easy-to-follow method to investigate people's perceptions and views. Unlike the Q, our protocol does not aim to compare patterns of opinions between stakeholder groups – hence, the absence of factor analysis - but rather is an input to improve the transparency of stated preference techniques. Armatas *et al.* (2014) applied it to select and define environmental attributes for consequent non-market valuation. Attributes were selected based on the Q-sorts (scoring), the resulting factor arrays and several quotes gathered during the exit interviews. The Q is used in particular when a large list of components needs to be downsized to a manageable number to be included in a stated preference survey. Forced trade-offs required by the Q are absent in Likert-type surveys, where a respondent can assign a high level of importance to all attributes. Another element that differentiates our method from the Q, is that Q is applied on individuals whereas we standardised our protocol to suit both groups and individuals. Several limitations of the Q made us develop the proposed protocol. First, we viewed the Q as too time-consuming. Secondly, bias can also enter into a Q study during the data analysis phase as the selection of a particular factor solution is also a methodological value judgment (Webler *et al.*, 2009). Thirdly, the Q sorting exercise can be seen as restrictive and participants may refuse to follow the pattern requested. We favoured the free choice of score allocation because it enables to obtain a final ranking with scores being different for nearly all characteristics, hence providing us with a first prioritisation of CE attributes. By taking a normative approach, it was useful to allow respondents to freely express their preferences associated with the NPHK characteristics, and not only ES. By asking the single question "S*hould there be more of this characteristic?",* we avoided forcing a Gauss curve on the rankings. Moreover, participants' responses to this 'satisfaction' question revealed what characteristic should be prioritised and for which action is required. This question was essential to select the attributes and to capture information to identify attribute-levels. It remains that people need to find the process interesting and simple, hence the choice of using a visual card deck (and polling software). To explore trade-offs more in-depth, a picture-based Q or a deliberative CE exercise, both applied in small groups could be seen as alternatives (Fairweather and R. Swaffield, 2001, Spash and Vatn, 2006, Spash, 2007).

Respondents' feedback about the card game was generally very positive. However, more background information about certain cards could have been provided. For instance, the card 'agricultural land' would need to be defined either as 'intensive' or 'extensive'. The

characteristic ‘natural forests’ was understood by some as primary or original forests, whereas we meant forests with original tree species. Also, ES unfamiliarity was a challenge despite the provision of clarification and definitions, in particular regarding ES such as ‘groundwater storage’ and ‘soil fertility’. Some cultural and regulating ES are abstract or intangible, and thus prove difficult to practically quantify and measure (van Zanten *et al.*, 2014). These observations emphasised the challenging roles of both the moderator who needs to have knowledge of the topic and the researcher who selects the items and the pictures. It is essential to have a clear description and knowledge of each characteristic prior to the application of the semi-qualitative protocol, because the level of knowledge regarding the type and quantity of benefits deriving from the park characteristics vary greatly between stakeholder groups. To ensure that all respondents understand the meaning of each characteristic, we suggest to include factual information on the cards; for instance for ‘drinking water’ to write that 15% of the water provided in the region (Flanders) flows from the NPHK. Another option to avoid misunderstandings or misinterpretations of ES and biodiversity features, could be to explain the ES concept individually to respondents prior to the FGs and INTs (Cárcamo *et al.*, 2014). Although misspecification of attributes is still possible using qualitative techniques, the rigorousness of the data collection process reduces that risk (Kløjgaard *et al.*, 2012).

Similar to discourse-based methods, the purpose of this study was to reach an agreement on what park characteristics are most important. For this, the proposed protocol requires that FGs are homogeneous in order to then add them to the INTs without introducing weights. Additionally, the protocol is standardised across the two qualitative methods to ensure that adding the results from both FGs and INTs is feasible.

The respondents represented a wide range of stakeholders; however, some stakeholder groups consisted of only one participant. A downside of qualitative approaches is that involving more than a small fraction of the citizens is very demanding. Thus, we must accept to be represented by someone (Vatn, 2009). As stated by O’Neill (2001) and Vatn (2009) we agree that statistical representation is not an issue and that it is rather about proposing a procedure by which representatives can act legitimately on behalf of others.

Another challenge often reported in social valuation studies is that people are able to freely rate a wide variety of ES. However, it is rarely the case that all ESs can be provided simultaneously, but respondents may be unaware of this during the moment of valuation (Scholte *et al.*, 2015). Consequently, it is the task of the experimental designer to ensure that what will be presented to participants is credible.

One has to bear in mind that the appropriate characterization of a service or value is dependent on appropriate methods, and no method is universally applicable (Chan *et al.*, 2012). Hence,

the question of how much or little qualitative work is needed before designing a CE is largely context-specific. The challenge is to construct choice tasks that enhance the CE's validity and quality. A mixture of qualitative methods, literature review and expert appraisal is therefore essential for the selection of stakeholder-, policy-, ecologically and economically relevant attributes and levels. Collaboration with experts (including among others economists and ecologists) is highly recommended to obtain a final combination of attributes and levels that do not overlap. We encourage future practitioners to use our protocol and adapt it to their study context, in particular if they aim to design a CE. Like with any qualitative approaches, the amount of time and money required to apply and adapt the entire protocol is dependent on numerous factors such as the researcher's skills, the stakeholders' profiles and readiness to participate, the aim of the study, the sensitivity of the topic and the natural resources at stake. The proposed protocol is, however, meant to be simple, short, and easily transferable. A scoring exercise followed by our 'satisfaction' question is easily applicable. Still, to test its full transferability it would need to be applied in other contexts with different stakeholders. Nevertheless, on the one hand literature tells us that in regions with a low literacy rate, picture-based processes are more adapted (Kenter *et al.*, 2011). Scales, on the other hand may be answered differently according to the local culture. Also, depending on the power dynamics of the group, methods may need to be employed that balance power between participants to avoid disregarding the views of the less powerful (Reed, 2008).

In conclusion, we would also like to note that, whereas our approach can be seen as semi-qualitative, its goal is to facilitate the integration of ES information into choice experiments. We do not advocate that the proposed protocol is a substitute for existing non-monetary valuation tools nor do we advocate that economic valuation is an alternative for non-monetary valuation.

7. CONCLUSION

This study proposes a semi-qualitative protocol for the selection of CE attributes and attribute-levels, and this in an environmental context. The process consists of 5 stages: 1) identification of stakeholders and characteristics of the study area, 2) creation of a card deck and a discussion protocol, 3) FGs and INTs, 4) semi-qualitative data analysis, and 5) selection of final attributes and attribute-levels for inclusion in a CE. This selection process was applied to the National Park Hoge Kempen in Belgium, and reveals the characteristics that are most important to a wide range of stakeholders.

This paper responds to the need for more rigorous CE attribute selection processes. It is the first study in the environmental research area to propose a detailed and easy-to-use selection process that includes the identification of both attributes and attribute-levels. The protocol presented in this paper is meant to be short, easily transferable, cost-effective, visually attractive and interesting to both participants and future CE practitioners. It is also particularly useful in contexts where many different stakeholders are involved. In an environmental or ES context, we advise to systematically apply a set of different methods and make use of visual tools. Also, in order to combine outcomes from both FGs and INTs, one needs to design a standardised protocol and to moderate homogeneous groups.

Future CE practitioners need to be aware that there is no standard rule for combining and reducing attributes (Louviere *et al.*, 2000). The semi-qualitative selection protocol described and applied here is not suggested as a magic bullet, but as an alternative to existing methods that have been applied in order to value natural resources. The study topic and the results the researcher aims to obtain from the CE undoubtedly influence the attribute selection. In our upcoming CE, we will calculate the willingness-to-pay of the park visitors for the five attributes selected in this study: biodiversity, chance to observe the Red deer, environmental education, air purification, and an entrance fee**.**

It is increasingly acknowledged that stakeholder engagement is an essential step to understand perceptions of the current and potential development scenarios, as well as to gain future support. This attribute selection approach is reported in a way that ensures its replicability in other contexts and that is suited for practitioners unfamiliar with qualitative research techniques and scientific research.

Acknowledgements

We would like to thank study informants, focus group participants and interviewees, as well as Tom Kuppens, Silvie Daniels, Janka Vanschoenwinkel and Michele Moretti of the Environmental Economics Research group of Hasselt University. Johan Van den Bosch, project leader at Regional Landschap Kempen en Maasland (RLKM), and Rolinde Demeyer from the Research Institute for Nature and Forest (INBO) have provided this study with valuable assistance and comments. Sarah Jeanloz was funded by the INTERREG IVB NEW program, Nele Witters and Sebastien Lizin are funded by Research Foundation- Flanders (FWO). Finally, we thank all reviewers for their constructive and insightful comments, and for their time.

References

Abiiro, G., G. Leppert, G. Mbera, P. Robyn and M. De Allegri (2014). "Developing attributes and attribute-levels for a discrete choice experiment on micro health insurance in rural Malawi." BMC Health Services Research **14**(1): 235.

Armatas, C. A., T. J. Venn and A. E. Watson (2014). "Applying Q-methodology to select and define attributes for non-market valuation: A case study from Northwest Wyoming, United States." Ecological Economics **107**(0): 447-456.

Barkmann, J., K. Glenk, A. Keil, C. Leemhuis, N. Dietrich, G. Gerold and R. Marggraf (2008). "Confronting unfamiliarity with ecosystem functions: The case for an ecosystem service approach to environmental valuation with stated preference methods." Ecological Economics **65**(1): 48-62.

Bateman, I. J. C., R. T.; Day, B.; Hanemann, M.; Hanley, N.; Hett, T.; Jones-Lee, M.; Loomes, G.; Mourato, S.; Özdemiroğlu, E.; Pearce, D. W.; Sugden, R.; Swanson, J. (2002). Economic valuation with stated preference techniques: a manual.

Bennett, J. and R. Blamey (2001). The Choice Modelling Approach to Environmental Valuation. Cheltenham, Edward Elgar Publishing.

Birol, E. and P. Koundouri (2008). Choice Experiments Informing Environmental Policy: A European Perspective, Edward Elgar Publishing, Incorporated.

Blamey, R. K., J. W. Bennett, J. J. Louviere, M. D. Morrison and J. C. Rolfe (2002). "Attribute Causality in Environmental Choice Modelling." Environmental and Resource Economics **23**(2): 167-186.

Brouwer, R., N. Powe, R. K. Turner, I. Bateman and I. H. Langford (1999). "Public attitudes to contigent valuation and public consultation." Environmental Values **8**: 325-347.

Cárcamo, P. F., R. Garay-Flühmann, F. A. Squeo and C. F. Gaymer (2014). "Using stakeholders' perspective of ecosystem services and biodiversity features to plan a marine protected area." Environmental Science & Policy **40**(0): 116-131.

Catt, H. (2002). Democracy in Practice, Taylor & Francis.

Cerda, C. and T. Losada (2013). "Assessing the value of species: a case study on the willingness to pay for species protection in Chile." Environ Monit Assess **185**(12): 10479-10493.

Cerda, C., A. Ponce and M. Zappi (2013). "Using choice experiments to understand public demand for the conservation of nature: A case study in a protected area of Chile." Journal for Nature Conservation **21**(3): 143-153.

Chaminuka, P., R. A. Groeneveld, A. O. Selomane and E. C. van Ierland (2012). "Tourist preferences for ecotourism in rural communities adjacent to Kruger National Park: A choice experiment approach." Tourism Management **33**(1): 168-176.

Chan, K. M. A., A. D. Guerry, P. Balvanera, S. Klain, T. Satterfield, X. Basurto, A. Bostrom, R. Chuenpagdee, R. Gould, B. S. Halpern, N. Hannahs, J. Levine, B. Norton, M. Ruckelshaus, R. Russell, J. Tam and U. Woodside (2012). "Where are Cultural and Social in Ecosystem Services? A Framework for Constructive Engagement." BioScience **62**(8): 744-756.

Coast, J., H. Al-Janabi, E. J. Sutton, S. A. Horrocks, A. J. Vosper, D. R. Swancutt and T. N. Flynn (2012). "Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations." Health Economics **21**(6): 730-741.

Coast, J. and S. Horrocks (2007). "Developing attributes and levels for discrete choice experiments using qualitative methods." Journal of Health Services Research & Policy **12**(1): 25-30.

Cowling, R. M., B. Egoh, A. T. Knight, P. J. O'Farrell, B. Reyers, M. Rouget, D. J. Roux, A. Welz and A. Wilhelm-Rechman (2008). "An operational model for mainstreaming ecosystem services for implementation." Proceedings of the National Academy of Sciences **105**(28): 9483-9488.

Darvill, R. and Z. Lindo (2015). "Quantifying and mapping ecosystem service use across stakeholder groups: Implications for conservation with priorities for cultural values." Ecosystem Services **13**(0): 153-161.

Dudley, N. E. (2008). Guidelines for Applying Protected Area Management Categories. Gland, World Conservation Union (IUCN).

Eagles, P. F. J., F. Romagosa, W. C. Buteau-Duitschaever, M. Havitz, T. D. Glover and B. McCutcheon (2013). "Good governance in protected areas: an evaluation of stakeholders' perceptions in British Columbia and Ontario Provincial Parks." Journal of Sustainable Tourism **21**(1): 60-79.

Fairweather, J. R. and S. R. Swaffield (2001). "Visitor Experiences of Kaikoura, New Zealand: an interpretative study using photographs of landscapes and Q method." Tourism Management **22**(3): 219-228.

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

Forrester, J., B. Cook, L. Bracken, S. Cinderby and A. Donaldson (2015). "Combining participatory mapping with Q-methodology to map stakeholder perceptions of complex environmental problems." Applied Geography **56**: 199-208.

Gómez-Baggethun, E. and E. Kelemen (2008). Participatory Methods for Valuing Ecosystem Services. THEMES Summer School, Lisbon.

Greiner, R., M. Bliemer and J. Ballweg (2014). "Design considerations of a choice experiment to estimate likely participation by north Australian pastoralists in contractual biodiversity conservation." Journal of Choice Modelling **10**(0): 34-45.

Hein, L., K. van Koppen, R. S. de Groot and E. C. van Ierland (2006). "Spatial scales, stakeholders and the valuation of ecosystem services." Ecological Economics **57**(2): 209-228.

Hiligsmann, M., C. van Durme, P. Geusens, B. G. C. Dellaert, C. D. Dirksen, T. van der Weijden, J.-Y. Reginster and A. Boonen (2013). "Nominal group technique to select attributes for discrete choice experiments: an example for drug treatment choice in osteoporosis." Patient preference and adherence **7**: 133-139.

Hoyos, D. (2010). "The state of the art of environmental valuation with discrete choice experiments." Ecological Economics **69**(8): 1595-1603.

Hoyos, D., P. Mariel, U. Pascual and I. Etxano (2012). "Valuing a Natura 2000 network site to inform land use options using a discrete choice experiment: An illustration from the Basque Country." Journal of Forest Economics **18**(4): 329-344.

Johnston, R. J., E. T. Schultz, K. Segerson, E. Y. Besedin and M. Ramachandran (2012). "Enhancing the content validity of stated preference valuation : the structure and function of ecological indicators." Land economics : a quarterly journal of planning, housing & public utilities **88**(1, (2)): 102-120.

Kaffashi, S., M. R. Yacob, M. S. Clark, A. Radam and M. F. Mamat (2015). "Exploring visitors' willingness to pay to generate revenues for managing the National Elephant Conservation Center in Malaysia." Forest Policy and Economics(0).

Kaplowitz, M. D. and J. P. Hoehn (2001). "Do focus groups and individual interviews reveal the same information for natural resource valuation?" Ecological Economics **36**(2): 237-247.

Kenter, J. O., T. Hyde, M. Christie and I. Fazey (2011). "The importance of deliberation in valuing ecosystem services in developing countries—Evidence from the Solomon Islands." Global Environmental Change **21**(2): 505-521.

Kløjgaard, M. E., M. Bech and R. Søgaard (2012). "Designing a Stated Choice Experiment: The Value of a Qualitative Process." Journal of Choice Modelling **5**(2): 1-18.

Krasny, M. E., A. Russ, K. G. Tidball and T. Elmqvist (2014). "Civic ecology practices: Participatory approaches to generating and measuring ecosystem services in cities." Ecosystem Services **7**: 177-186.

Krueger, R. A. and M. A. Casey (2000). Focus Groups. A Practical Guide for Applied Research (3rd Edition), Thousand Oaks, CA: Sage Publications.

Lancaster, K. (1966). "A new approach to consumer theory." Journal of Political Economy **74**: 132-157.

Lancsar, E. and J. Louviere (2006). "Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences?" Health Economics **15**(8): 797-811.

Li, C.-z., J. Kuuluvainen, E. Pouta, M. Rekola and O. Tahvonen (2004). "Using Choice Experiments to Value the Natura 2000 Nature Conservation Programs in Finland." Environmental and Resource Economics **29**(3): 361-374.

Liu, S., R. Costanza, S. Farber and A. Troy (2010). Valuing ecosystem services Theory, practice, and the need for a transdisciplinary synthesis. Ecological Economics Reviews. K. Limburg and R. Costanza. **1185:** 54-78.

Liu, X. and K. W. Wirtz (2010). "Managing coastal area resources by stated choice experiments." Estuarine Coastal and Shelf Science **86**(3): 512-517.

Lo, A. Y. (2013). "Agreeing to pay under value disagreement: Reconceptualizing preference transformation in terms of pluralism with evidence from small-group deliberations on climate change." Ecological Economics **87**: 84-94.

Louviere, J. J., D. A. Hensher and J. D. Swait (2000). Stated Choice Methods: Analysis and Applications. Cambridge University Press.

Martín-López, B., C. Montes and J. Benayas (2007). "The non-economic motives behind the willingness to pay for biodiversity conservation." Biological conservation **139**(1): 67-82.

Maynard, S., D. James and A. Davidson (2014). "Determining the value of multiple ecosystem services in terms of community wellbeing: Who should be the valuing agent?" Ecological Economics(0).

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. New York, Frontiers in Econometrics Academic Press.

Michaels-Igbokwe, C., M. Lagarde, J. Cairns and F. Terris-Prestholt (2014). "Using decision mapping to inform the development of a stated choice survey to elicit youth preferences for sexual and reproductive health and HIV services in rural Malawi." Social Science & Medicine **105**(0): 93-102.

Millennium Ecosystem Assessment (2005). Ecosystems and Human Well-being: Synthesis. Washington, DC.

Petheram, L. and B. M. Campbell (2010). "Listening to locals on payments for environmental services." Journal of Environmental Management **91**(5): 1139-1149.

Rajmis, S., J. Barkmann and R. Marggraf (2009). "User community preferences for climate change mitigation and adaptation measures around Hainich National Park, Germany." Climate Research **40**: 61-73.

Reed, M. S. (2008). "Stakeholder participation for environmental management: A literature review." Biological Conservation **141**(10): 2417-2431.

Rolfe, J., -, K. Alam, J. Windle and S. M. Whitten (2004). Designing the Choice Modelling Survey Instrument for establishing riparian buffers in the Fitzroy Basin.

Rolfe, J., J. Bennett and J. Louviere (2000). "Choice modelling and its potential application to tropical rainforest preservation." Ecological Economics **35**(2): 289-302.

Rolfe, J. and J. Windle (2015). "Testing attribute selection and variation in a choice experiment to assess the tradeoffs associated with increased mining development." Land Use Policy **42**(0): 673-682.

Ryan, M., V. Watson and V. Entwistle (2009). "Rationalising the 'irrational': a think aloud study of discrete choice experiment responses." Health Economics **18**(3): 321-336.

Scholte, S. S. K., A. J. A. van Teeffelen and P. H. Verburg (2015). "Integrating socio-cultural perspectives into ecosystem service valuation: A review of concepts and methods." Ecological Economics **114**: 67-78.

Shoyama, K., S. Managi and Y. Yamagata (2013). "Public preferences for biodiversity conservation and climate-change mitigation: A choice experiment using ecosystem services indicators." Land Use Policy **34**: 282-293.

Spash, C. L. (2002). "Informing and forming preferences in environmental valuation: Coral reef biodiversity." Journal of Economic Psychology **23**(5): 665-687.

Spash, C. L. (2007). "Deliberative monetary valuation (DMV): Issues in combining economic and political processes to value environmental change." Ecological Economics **63**(4): 690-699.

Spash, C. L. and A. Vatn (2006). "Transferring environmental value estimates: Issues and alternatives." Ecological Economics **60**(2): 379-388.

Swedeen, P. (2006). "Post-normal science in practice: A Q study of the potential for sustainable forestry in Washington State, USA." Ecological Economics **57**(2): 190-208.

van Zanten, B., P. Verburg, M. Espinosa, S. Gomez-y-Paloma, G. Galimberti, J. Kantelhardt, M. Kapfer, M. Lefebvre, R. Manrique, A. Piorr, M. Raggi, L. Schaller, S. Targetti, I. Zasada and D. Viaggi (2014). "European agricultural landscapes, common agricultural policy and ecosystem services: a review." Agronomy for Sustainable Development **34**(2): 309-325.

Vatn, A. (2009). "An institutional analysis of methods for environmental appraisal." Ecological economics **68**(8-9): 2207-2215.

Wang, P.-W. and J.-B. Jia (2012). "Tourists' willingness to pay for biodiversity conservation and environment protection, Dalai Lake protected area: Implications for entrance fee and sustainable management." Ocean & Coastal Management **62**(0): 24-33.

Webler, T., Danielson, S., & Tuler, S. (2009). Using Q method to reveal social perspectives in environmental research. Greenfield MA: Social and Environmental Research Institute. Downloaded from: http://www.seri-us.org/sites/default/files/Qprimer.pdf

Wilson, M. A. and R. B. Howarth (2002). "Discourse-based valuation of ecosystem services: establishing fair outcomes through group deliberation." Ecological Economics **41**(3): 431-443.

Zander, K. K. and A. G. Drucker (2008). "Conserving what's important: Using choice model scenarios to value local cattle breeds in East Africa." Ecological Economics **68**(1-2): 34-45.

APPENDIX A

FOCUS GROUP and INTERVIEW PROTOCOL

Generic version

Project Name:

Last Updated:

Project Manager:

MODERATING FOCUS GROUPS

The moderating team is composed of:

1 moderator

2 assistant moderators

The **moderator** is a source of data! He or she is primarily concerned with directing the discussion, keeping the conversation flowing, and taking a few notes when possible.

As a moderator, it is important to present yourself as a researcher rather than a friend. You will need to let participants know that you are part of a team that is conducting research for the region's development. This formality communicates to participants that their participation is important.

The moderator should possess the following skills:

- Understand what the goals of the projects are, who the participants are, and be familiar with this guide
- Can listen attentively. Are participants answering the question?
- Know when to probe for more information ("Is there anything else?") and when to move on
- Ability to listen and think at the same time
- Adequate knowledge of the topic
- Can keep personal views and ego out of the facilitation; avoid giving personal opinions
- Can appropriately manage challenging group dynamics

The **assistant moderators** take comprehensive handles the environmental conditions and logistics, and respond to unexpected interruptions. The assistant is key during the post-meeting analysis of the session.

Assistant moderator n°1:

- Runs a tape recorder during the session
- Runs the PowerPoint presentation(s) (*e.g.* pictures shown to participants + TurningPoint)

Assistant moderator n°2:

- Takes notes: people's reactions and the way they interact with each other

After each group discussion, the moderator and assistant moderator(s) need to contribute their observations in a debriefing session (with the researcher – research team).

THE PROTOCOL

STAGE 1

Identify park characteristics from the literature and consultations with experts (*e.g.* researchers, natural resource managers)

Identify key stakeholders via purposive sampling (*e.g.* internet-based search, snowball sampling)

STAGE 2

Create a picture-card deck representing the pre-selected characteristics and the categories (*e.g.* 1 = very important to 5 = not at all important)

STAGE 3

Step-by-step protocol during FGs and INTs

Step	Focus group	Interview
1. Introduction *10 min.*	• Brief presentation of the study • Explain why the participants' opinions are important • Explain the "rules of the game"	
2. Scoring *25-30 min.*	• Use of the TurningPoint software: each participant has a voting remote control • Display the picture slide show • Participants attribute a score to each picture, from 1 (very important) to 5 (not important at all)	• The respondent holds the picture-card deck • The respondent places cards one by one onto the five categories (1 to 5)
3. Ranking *10 min.*	• Individual scoring is averaged to obtain the ranking of the group (preferable to do that during a break) • Display the top 15 to 20 cards of the ranking (Excel table)*	Consider only the characteristics categorised as most important (1) (and important (2) in case very few cards were placed in category 1)
4. Discussion about ranking *10 min.*	• Participants can make changes and add new characteristics to the list • Participants reach a (broad) consensus about the top characteristics	• The respondent can make changes to the list and add new characteristics to the list • The respondent shall be satisfied with his final selection
5. Satisfaction with the current situation *15-20 min.*	• Consider the top characteristics one by one • Participants answer to the question *'Should there be more of this characteristic?'* • Report Yes/No/Not applicable/50-50 (if opinions are divided)	• Consider the top characteristics one by one • The respondent answers to the question *'Should there be more of this characteristic?'* • Report Yes/No/Not applicable
6. Closure *10 min.*	• Ask to fill out the consent form • Thank • Mention a potential future feedback	

* Example of a table composed of the top characteristics presented to a stakeholder group

Rank	Characteristic	Mean score	*Should there be more?*		
			Yes	*No*	*N.A.*
1	Biodiversity refuge	1.10	x		
2	Natural forests	1.10	x		
3	Plant and animal conservation	1.10	x		
4	Heathlands	1.20		x	
5	Wetlands	1.20		x	
6	Peace and quiet	1.20	x		
7	Air purification	1.20	x		
8	Landscape variety	1.30	x	x	
9	Carbon storage	1.40	x	x	
10	Drinking water	1.40			x
11	Walking	1.50		x	
12	Pollination	1.50	x		
13	Restoration of disturbed habitats	1.50	x		
14	Groundwater storage	1.60	x		
15	Education	1.60		x	

STAGE 4

- ***Analysis of the scoring exercise***

Aggregate top list: take the average scores of each FG and INT and calculate the total average score for all characteristics

Test differences among groups using a non-parametric test (*e.g.* Kruskal Wallis) *(optional)*

- ***Analysis of the interactive discussion***

Recode answers to the 'satisfaction' question:

Yes = 2, 50/50 = 1, No or Not applicable = 0

Sum up the numbers 0,1,2; the total of points that can be obtained is 2 x the number of stakeholder groups (in our case: 9 x 2 = 18 points)

If the sum of the of points for a characteristic is higher than the half (in our case: > 9), it means that the majority of stakeholders would like to see a change for this characteristic
→ Your CE candidates meet two conditions: have the highest scores and obtain a number of points that is over half of the mean number

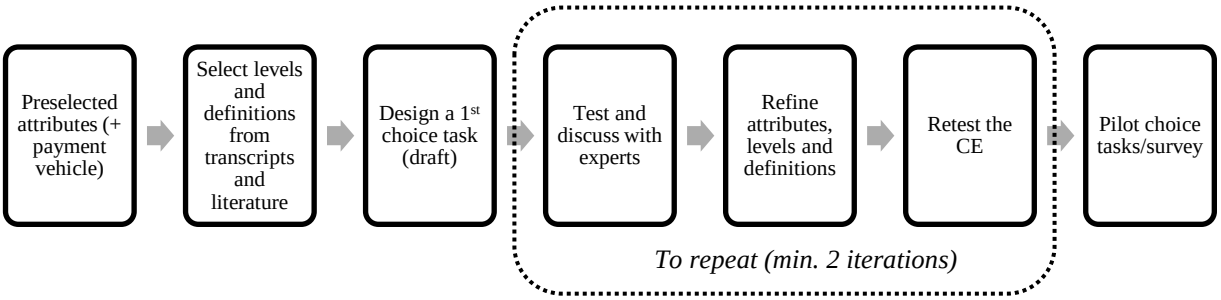
Transcribe INT tapes (partially) / use FG notes as abridged transcripts: to be used for the final selection and wording of CE attributes

STAGE 5

The attribute selection process is guided by the **scores**, the **answers to the 'satisfaction' question**, the **arguments** raised in the interactive discussion, and the following **inclusion-exclusion criteria**, whereby attributes should be:

- not overlapping other attributes (to avoid inter-attribute correlation) / mutually exclusive in nature
- policy-relevant
- demand-relevant: important, understandable and meaningful to people and relate to their reasons for having willingness-to-pay to conserve biodiversity and ES
- measurable
- ecologically and economically relevant
- describable by combining simple explanations and visual instruments such as photographs, charts, and pictures
- limited to a number between 4 and 8 because trade-offs become difficult to understand

Once you have the preselected attributes, apply the following steps to select the final attributes and design the pilot choice tasks:



Note: we used Ngene software to design the choice tasks