Chapter 2
Developing the Questionnaire

Abstract This chapter outlines the essential topics for developing and testing a questionnaire for a discrete choice experiment survey. It addresses issues such as the description of the environmental good, pretesting of the survey, incentive compatibility, consequentiality or mitigation of hypothetical bias. For the latter, cheap talk scripts, opt-out reminders or an oath script are discussed. Moreover, the use of instructional choice sets, the identification of protest responses and strategic bidders are considered. Finally, issues related to the payment vehicle and the cost vector design are the subject of this section.

2.1 Structure of the Questionnaire

According to Dillman et al. (2008), a good questionnaire is like a conversation that has a clear, logical order. This includes to begin with easily understandable, salient questions and grouping-related questions with similar topics. Especially in web surveys, the initial questions have to be chosen carefully. Respondents cannot have a look at all the survey questions as with mail surveys and, therefore, the initial questions are crucial to get them interested in the survey. These questions should therefore apply to all respondents. Also, in the introduction to the survey respondents should be informed about the topic of the survey and give their consent to participate. There is evidence that an interesting survey topic can increase the response rate (Groves et al. 2004; Zillmann et al. 2014) and this can be taken into account in the introduction to the survey. While it is difficult to estimate a topic-related selection bias in survey participation, researchers should consider such a potential bias (e.g. Nielsen et al. 2016). For instance, it is more likely that a survey on environmental issues might be answered by individuals who are interested in environmental issues or have a high level of environmental concern. Such a potential bias could be reduced by making the survey and survey topic more general (e.g. quality of life in a region which also includes environmental issues).

© The Author(s) 2021

P. Mariel et al., *Environmental Valuation with Discrete Choice Experiments*, SpringerBriefs in Economics, https://doi.org/10.1007/978-3-030-62669-3_2

In some surveys, respondents have to be screened out at the beginning of the survey because they do not belong to the target group. In this case, both eligible and ineligible respondents should be directed to the main survey after answering the screening questions in order to record non-response. Those who are ineligible should receive a thank you statement after being screened out.

It is a well-established fact that responses to survey questions can be affected by question context (Schuman et al. 1981; Tourangeau et al. 2000; Moore 2002; Dillman et al. 2008). Two types of context effects can be distinguished (Tourangeau et al. 2000, p. 198). First, a directional context effect is present if answers to a target question such as choice experiment tasks depend on whether context questions such as relevant attitudinal questions are placed before or after the target question. Second, a correlational context effect occurs if the correlation between responses to the target and the context questions is affected by the question order. The latter means, for example, that the relationship between attitude measurements and responses to choice tasks is affected by question order. Question context is likely to affect stated preferences because surveying relevant attitudes prior to choice tasks might provide an "interpretive framework" (Tourangeau and Rasinski 1988) with regard to the choice questions, leading to possible judgement effects (Tourangeau and Rasinski 1988, p. 306). There are only a few studies which have tested this type of context effects in SP surveys. Pouta (2004) showed in a contingent valuation study that the inclusion of relevant belief and attitudinal questions prior to the valuation question increases the likelihood that an environmentally friendly alternative is chosen and increases the respondents' WTP for environment forest regeneration practices in Finland. Liebe et al. (2016) find positive evidence for a directional context effect in a choice experiment study on ethical consumption. Therefore, when constructing a questionnaire it is important to be aware of this and consider possible implications of the fact that stated preferences and corresponding WTP estimates are likely to be affected by whether relevant attitudes are surveyed before or after the choice tasks in the experiment. In some cases, it may be considered relevant to ensure that respondents have thought about their own attitudes before answering the preference eliciting choice tasks, in other cases not.

Since respondents should be able to make informed decisions in line with their interests, the hypothetical market has to be described in as much detail as possible. This does not mean overloading respondents with information but naming the most important characteristics of the market context. Table 2.1 gives an overview of these characteristics (see Carson 2000, p. 1415 for contingent valuation) as well as a structure of a typical choice experiment questionnaire for environmental valuation.

When asking for preferences of unfamiliar goods or services, researchers might want to place questions on attitudes, social norms, etc., prior to the choice tasks in order to make respondents think carefully about the topic before answering the choice questions (see Bateman et al. 2002, p. 150, who recommend asking attitudinal and opinion questions before the valuation section in contingent valuation surveys). On the other hand, the literature on context effects discussed above (e.g. Liebe et al. 2016) often suggests asking such questions after the choice tasks instead because answering questions which are relevant for the choice task might activate socially

desirable response behaviour or direct attention to specific choice attributes, which is probably unintended by the researcher. Therefore, researchers should consider the possibility of unintended context effects which can also apply to so-called warm-up questions or instructional choice sets before the actual choice tasks.

Socio-demographic questions including gender, age, education and income are generally asked at the end of the survey. This is typically recommended because they refer to personal and partly sensitive information. The income question is especially sensitive and often causes high item non-response and missing values. On the other hand, it is an important variable for economic valuation studies. One way to reduce item non-response is to first ask respondents for an exact income amount and in the event they refuse to answer or choose a do not know option, provide a list of income

Table 2.1 Structure of a typical DCE survey

Questionnaire section	Content/related questions
Information given in the introduction to the survey	• What is the aim of the survey? • Who is eligible to take part? • Who is conducting the survey? • To what extent is anonymity or confidentiality of survey responses guaranteed? • How are the results going to be used/disseminated? • Initial question(s) which are easy to understand and applicable to all respondents • How long will it take to answer the survey? • Is ethical approval obtained for the study? • How will the data be stored?
Behavioural questions	• How often have you visited the environmental good in question (e.g. a forest or a beach)? • Which activities did you undertake?
Introduction in which the context of the choice task is described	• What is the societal or environmental problem? • How is the environmental good at hand linked to it?
Detailed description of the environmental good at hand, the institutional setting and payment vehicle	• What are the attributes of the good and how can they vary? • Who is responsible for the provision of the good (public or private institution, etc.)? • How do respondents pay for the good (taxes, fees, contributions to a fund, etc.)?
Choice experiment tasks	• How many alternatives shall a choice task include? • When shall a status quo and/or opt-out option be included? • How many choice tasks shall be presented to a respondent?

(continued)

Table 2.1 (continued)

Questionnaire section	Content/related questions
Follow-up questions to the choice tasks	• How did the respondents make their choices? • What were the most important choice attributes? • How difficult was it to answer the choice tasks? • To what extent is protesting affecting choice behaviour?
Questions on relevant attitudes, norms, etc., which help to "explain" heterogeneity of stated preferences	• To what extent are the respondents in favour or disfavour of the environmental good at hand (i.e. specific attitudes)? • How much are they concerned about environmental protection (i.e. general attitudes)? • To what extent does the social environment reward paying for an environmental good (i.e. social norm)? • To what extent does a respondent perceive a moral obligation to pay for the good at hand (i.e. personal norm)?
Questions on the socio-demographic background	• What is the respondent's gender, age, education, income, etc.?

categories (Duncan and Petersen 2001). Alternatively, income bands can be used to increase the response rate to the income variable. Only if the study is based on a quota design socio-demographic questions (often gender, age and education) are typically asked at the beginning of the survey to control sample quota and to screen out respondents in case of filled sampling quotas.

Another aspect of questionnaire design that can influence survey participation and dropouts is the length of a questionnaire (Galesic and Bosnjak 2009). Typically, this also depends on the survey mode—face-to-face interviews can be longer than mail and web surveys (see Sect. 4.2). Clearly, shorter surveys (e.g. around 20 min) are preferred over longer surveys. For example, regarding web surveys it can be shown that the longer the stated survey length in the introduction to the survey the lower the likelihood of participating and completing the survey (Galesic and Bosnjak 2009). Likewise, it is often expected that difficult questions cause higher dropout rates. Furthermore, in longer questionnaires the answers to questions positioned later in the questionnaire can be less valid compared to positioning the same questions at the beginning of the questionnaire. It is also important to state a reliable survey length in the introduction to the survey.

In summary, researchers need to be aware that all aspects of the questionnaire can affect the results of a DCE survey. A well-designed introduction to the survey can reduce non-response (unwillingness to participate) and selection (participation of specific individuals) bias. The question order can have an influence on choice experiment results. It is therefore important to consider unintended context effects,

for example, that environmental attitudes surveyed before the choice tasks might affect responses to the choice tasks, which might or might not be in the interest of the researchers. Also, warm-up questions or instructional choice sets can cause unintended anchoring effects, starting point effects, etc. While the ("optimal") length of a questionnaire also depends on the survey mode, it can be recommended to aim for shorter questionnaires (e.g. around 20 min), which not only increase survey participation, but also positively affect the validity and reliability of survey responses. The books by Dillman et al. (2008, 2014) are recommended as a comprehensive and detailed introduction to survey and questionnaire design.

2.2 Description of the Environmental Good

In addition to generic issues regarding how ordering may affect the answers to other questions (Sect. 2.1) another central issue is how much, and which information, to provide to respondents before presenting the choice sets. The basic principle is that a clear, unambiguous description (including time, scope, etc.) of the good to be valued is always required.

In many environmental valuation settings, valuation is conducted for a specific policy, resulting in a marginal change in the provision of certain goods. Describing the policy is therefore an important part of setting the context of the hypothetical market. The researcher typically wants respondents to perform trade-offs in a specific situation—e.g. evaluating policy proposals or choices of recreational actions. In order to make such choices, respondents need to be informed on what the choice is about. This involves explaining the *policy context* (the overall aim), the *environmental consequences* it will have (as also explained by the attributes) and how the *hypothetical market* is set up (e.g. how payment is to be made).

In principle, we want respondents to represent the target population, i.e. we should not provide any information at all if they also make uninformed decisions in real life—as it is well known that information affects choices (Jacobsen et al. 2008). On the other hand, we would like people to make informed choices: if the results should inform policymakers, they should reflect the preferences of people and most people would be likely to seek information before making choices. But the level of information seeked may vary widely as we already know from real referenda.

A good starting point for deciding how and to what extent respondents should be informed is to think about the amount and quality of information people might already have prior to the survey. In the case of a local good, many people might already be familiar with it and also with its present quality and if so, little information may be enough. In case of a unfamiliar endangered species to be protected on a different continent obviously this might be different, suggesting that more information is needed. The risk of too much information is biasing people, the risk of too little is that if respondents do not have sufficient information on the good, they may use their imagination and hence different respondents end up valuing different "goods". Often we are in situations where we would like to provide a lot of information to

respondents for them to make informed choices. But how much is enough? There is no clear answer to that. Pre-tests and focus groups can help to clarify this. The more unfamiliar a good and the less tangible it is, the more information is needed for them to make choices concerning a specific change. It may also differ depending on people's previous knowledge. A few examples may illustrate this: (1) working with farmers' willingness to change practices typically requires little information about the goods as they are well aware of their management practices and what they obtain from them, on the other hand, they may require more information on the instruments by which the practices have to be changed (e.g. Vedel et al. 2015a, b). Especially if we are working in developing countries with weak institutional settings (Nielsen et al. 2014; Rakotonarivo et al. 2017) it may require some effort to describe the hypothetical market (see, e.g., Kassahun and Jacobsen 2015). Working with recreational preferences in Western Europe often requires little information about the good and the hypothetical market, typically information on distance, as do preferences for environmental characteristics of food choices. Working with unfamiliar nature like deep water coral reefs requires a lot of information as many people have never heard about them (Aanesen et al. 2015).

A challenge with providing information is to ensure that people read it and digest it. Especially with online surveys, this is a big problem. Therefore, we often see that information is interrupted with attitudinal questions or questions about people's knowledge even though this may lead to context effects as mentioned in Sect. 2.1. For example, the description of the extent of a specific nature area may be accompanied with a question about whether people have visited the area or know the characteristics described. This may make them think specifically about this area and not nature areas in general (an intentional directional context, also referred to as framing), but can potentially bias them in terms of making them think more about recreational values than existence values. These kinds of trade-offs are important to consider and to test in focus groups and pre-survey interviews (see Sect. 2.3).

In the description explaining the environmental good, the hypothetical market and the policy situation, it is important to make the following points clear to respondents:

(a) That the proposed policy change leads to a certain outcome and that there is at least some scientific evidence for this relationship. A few examples may be: setting aside forests as a means of increasing the likelihood of securing the survival of endangered species; afforestation as a means of achieving greater carbon sequestration than the alternative land use under consideration, implementing restrictions on fertilisers in agriculture to affect water quality in nearby streams, etc. Notice that the relationship needs to be described as objectively as possible—both for validity and also to ensure that respondents are not protesting because they do not believe the given stated consequences. A particular challenge here is that the precise and objective description of these often quite complex biological relationships also has to be conveyed in layman's terms to be understandable to all respondents. In most cases, this requires careful testing in consecutive focus group interviews.

(b) That it is generally important to distinguish means from outcomes, and most often we do this after valuing the outcome (as the means can be assessed as a cost). But a challenge occurs if the proposed means to achieve a particular outcome has positive/negative side effects, e.g. creation/destruction of local jobs, or regulating invasive species by "inhuman" means. These are important to identify in focus groups and through interviews with experts, and if present, attempt to avoid them in the description, or use a specific attribute to eliminate the effect on other attributes—even if this attribute in itself is of little interest.
(c) That the proposed policy change leading to a particular outcome is perceived as realistic by respondents. Quite often describing the scientific basis (as mentioned in point a) can be challenging. It is also important that the relationship of what is being valued is related to aspects that matter to the respondent. A classic example is valuing water quality, where a possible measure is N-concentration. Nevertheless, to relate it to a value that matters to people it has to be translated into the final ecosystem services being provided which are those presumably affecting people's utility, e.g. clarity of water for swimming, effect on biodiversity, etc. (see, e.g., Jensen et al. 2019).
(d) That the ones described to carry out the policy also have the power to do so—i.e. that the institutional setting is realistic.
(e) That the scope of the change is made explicit. In the contingent valuation literature this has been strongly emphasised. In DCE, it has often drawn less attention as the attributes vary and thereby internal scope sensitivity is ensured. But to make sure that respondents understand this well, it is necessary to be quite specific about the scope of the project/project combination proposed.
(f) That the attribute and attribute levels are well defined and explained in an understandable setting.
(g) That attributes should vary independently from each other if possible. This can often be a problem in, for example, conservation, where endangered species conservation and habitat restoration are correlated. An example of distinguishing these is Jacobsen et al. (2008) who has an attribute related to the area conserved, and then an attribute of ensuring survival of endangered species. For this to be realistic to respondents (and according to the natural science basis), respondents were told that it would be possible because other management initiatives targeted endangered species. This may in fact be possible as illustrated in a conservation strategy paper using the same valuation data as input (Strange et al. 2007).
(h) That the payment vehicle is well described (see Sect. 2.10).
(i) That consequentiality is ensured (see Sect. 2.5) and as far as possible also incentive compatibility (see Sect. 2.4).

In conclusion, content validity (see Sect. 8.1) requires a precise description of the environmental consequences, the policy to be implemented and the hypothetical market in a sufficiently detailed way so that respondents can make informed choices. This has to be weighted carefully with the risk of biasing people if information is

not objective, or if a certain aspect of the good is emphasised over others. Thorough focus group and pilot testing are essential tools to find this balance.

Finally, another important issue not really touched on here is how information is conveyed to respondents. Most often information is provided through text and sometimes accompanying pictograms or images. Recently, several studies have started to use other media, for example, virtual environments (Bateman et al. 2009; Matthews et al. 2017; Patterson et al. 2017; Rid et al. 2018) or videos (Sandorf et al. 2016; Lim et al. 2020; Rossetti and Hurtubia 2020). To investigate whether those formats are more suitable to inform respondents about the good in question and the organisation of the hypothetical market is an open question requiring further research.

2.3 Survey Pretesting: Focus Groups and Pilot Testing

The development of SP surveys, as with all primary data-collection methods, requires devoting a substantial part of the overall work to designing and testing. Often this will be an iterative process that should use, among others, face-to-face pilot testing. Much effort should be devoted to translating expert knowledge into understandable and valuable information for respondents. Previous scientific investigation on the environmental characteristics of the good or service under valuation, expert advice and focus groups may facilitate the definition of attributes and levels of provision (Hoyos 2010). In this context, survey pretesting emerges as a basic prerequisite for a proper survey design (Mitchell and Carson 1989; Arrow et al. 1993; Johnston et al. 2017), including both qualitative (personal interviews or focus groups) and quantitative (pilot studies) pretesting. The main purpose of pretesting the survey is ensuring that the information provided in the questionnaire is sufficient, understandable and credible to the population, acknowledging that they may have different education levels and backgrounds. It is especially important to check that the environmental change, policy situation and hypothetical market (i.e. those highlighted in the previous section) are clear to respondents. Pretesting the survey also helps ensure the content validity of the questionnaire, as will be discussed in Sect. 8.1.

Testing the survey questionnaire generally involves four different methods: focus groups, cognitive interviews, group administrations and pilot surveys. Focus groups are small group (6–12 individuals) semi-structured, open-ended discussions among the relevant population. They facilitate the discussion of the concepts and language presented in the questionnaire and they are specifically useful in clarifying scenario and alternatives description, as well as evaluating the adequacy of the amount and level of information that respondents require in order to answer the valuation questions. Focus groups may also help when deciding the best strategy for explaining the task of making successive choices from a series of choice sets. Cognitive interviewing refers to questioning single individuals about his or her understanding and reactions to the questionnaire. Typically, concurrent verbal protocols are elicited from individuals in order to assess their understanding and reaction to the questionnaire. These protocols are especially useful to analyse respondents' reactions to specific sets of

the text using their own words (Willis 2005). Group administrations are designed for larger groups of people to silently record their answers to the questionnaire read to them by a professional interviewer (Wright and Marsden 2010).

Finally, pilot surveys are small field data collection testing with a small sample of the population (usually 50–100 respondents). They are highly recommended in order to develop a DCE survey because (1) it is cost-efficient as it may help detect problems in the questionnaire before collecting the whole sample; (2) it may serve as preliminary statistical analysis of the data; and (3) it may also help with defining priors for an efficient experiment design, as will be discussed in Sect. 3.2 (Leeuw et al. 2008). It is important to plan how participants for these different formats are recruited. While it is often convenient to use, for example, students, the question is whether participants who are easy to recruit sufficiently reflect the target population of the main survey. Good pretesting requires that people from the target population are involved in the pretesting phases.

Feedback from the respondents should be iteratively used in the revision of the questionnaire. Number of attributes and levels, payment vehicle and duration should be chosen in consonance with the good under valuation and its context. The analyst should weigh up the relevant number of attributes and the complexity of the design. The trade-off between the possibility of omission of relevant attributes and task complexity and cognitive burden to respondents may be analysed in focus groups and pilot surveys. Additionally, it may be interesting to use pre-tests to identify any possible interaction effect between attributes. Complexity of the choice task can be investigated with verbal protocols (Schkade and Payne 1994). In order to avoid group effects (Chilton and Hutchinson 1999), one-on-one interviews are also highly recommended (Kaplowitz and Hoehn 2001). With eye-tracking and other biometric sensor technology becoming increasingly affordable, it may be beneficial to supplement cognitive one-on-one interviews with such measures in order to acquire even more feedback on how respondents react to the information and questions presented to them in a questionnaire.

There is no fixed number of pre-tests of the survey that should be carried out because it may depend on the purpose of the study, the unfamiliarity of the good to be valued and the relative success of previous iterations, but current best practice recommends a minimum of four to six focus groups (Johnston et al. 2017). In cases like the BP Deepwater Horizon oil spill, for example, up to 12 focus groups, 24 cognitive interviews, 8 group administrations and 5 pilot surveys were conducted between mid-2010 and the end of 2013 for pretesting the questionnaire (Bishop et al. 2017). This is however an extreme case probably reflecting the largest amount of pretesting conducted in any SP survey. While the amount of pretesting needed is inherently case-specific and depends on the purpose for acquiring value estimates, for most environmental DCEs around 2–8 focus groups, 5–10 cognitive one-on-one interviews, and 1–2 pilot surveys would be considered sufficient.

Practitioners should bear in mind that proper pretesting of the survey requires time and, especially, resources for recruiting or rewarding participants, so that a specific budget for this purpose should be made available. It is also important to denote that gathering a random group from the relevant population may require the

pretesting of the survey in different locations when the market size is large (e.g. a nationwide survey). Some guides to methods of collecting data for testing the questionnaire include Morgan (1997), Krueger and Casey (2008) and Dillman et al. (2008). Finally, survey pretesting should be properly documented and made available for reviewing purposes.

2.4 Incentive Compatibility

Incentive compatibility is the process in which a truthful response to a question constitutes the optimal strategy for an agent (Carson and Groves 2007). This means that respondents should find it in their best interest to answer truthfully. And by construction this is problematic for hypothetical choices—because will it ever have an impact what respondents answer? If I am asked whether I would prefer to die in a car accident or from cancer, it is not incentive compatible: my answer will not affect my probability of dying from either. Nor is it a choice I will be in a situation to make. Therefore, I have no incentive to answer honestly. And when respondents do not have an incentive to answer honestly, we are not guaranteed to get honest answers reflecting the respondent's true preferences. Even worse, if they have an incentive to answer dishonestly (e.g. due to warm-glow giving), we may get very wrong answers. Incentive compatibility is found to be important in many empirical settings, and Zawojska and Czajkowski (2017) find in a review that when choices are incentive compatible, they are more likely to pass external validity tests.

To ensure incentive compatibility, Vossler et al. (2012) list the following requirements: (1) participants care about the outcome (see also Sect. 2.5), (2) payment is coercive—it can be enforced on everyone (see also Sect. 2.10), (3) a single binary (yes/no) question format is used, (4) the probability of project implementation is weakly monotonically increasing by the proportion of yes-voters. DCEs with more than a single choice set violate requirement number 3 and hence do not satisfy incentive compatibility conditions. Given that DCEs typically do not live up to the criteria of incentive compatibility, the question is how important it is? Therefore, various attempts can be made to investigate the importance of incentive compatibility.

One is to construct a provision rule: a mechanism can be constructed that ensures implementation of only one strategy and independence between choice sets. The latter is typically addressed in stated DCE by explicitly asking respondents to value the choice sets independently from each other.

Another possibility is to rely on only binary choice which can also be done in DCE (e.g. Jacobsen et al. 2008), but in such cases, only the first choice set is potentially incentive compatible, whereby little information is obtained from each individual. An approximation that is sometimes being used to ensure incentive compatible DCEs in experimental settings is to implement a premium mechanism that randomly draws one of the choice sets as the winner, and that policy is then implemented. In a setting of provision of a public good on a large scale this is problematic in practice and incentivised choices may be used instead. This is preferably related to the good in

question, but may also simply be a premium. Svenningsen (2019) is an example of the former. The incentive was formulated at several places throughout the survey. In the beginning as:

> The survey you are participating in now is a bit different than the usual survey. As mentioned in the invitation-email you are given the opportunity to earn up to 18,000 extra points, the equivalent of 200 DKK, by participating in this survey. During the survey you can choose to donate all 200 DKK or some amount below the 200 DKK/18,000 points to climate policy. More information on this will follow later in the survey.

Then before the choice sets (and split up on several screens):

> As mentioned in the invitation-email, you are given 200 DKK, the equivalent of 18,000 points extra for your participation in this survey. In the choices you are about to make you are free to spend some part of or all 200 DKK as a donation towards implementing the climate policy you choose. You are free to choose the amount you wish to keep, as well as the amount you wish to donate towards climate policy. The amount not spent in this survey will be transferred to your account with Userneeds before the 18th of March 2016. You will be asked to make 16 choices and in each of these choices you have to imagine that you have the full 200 DKK/18,000 points that you either can donate or keep. One of your choices will be drawn at random and paid out and you will be informed about which choice it was at the end of the survey. The choices from all participants will be added up and the total amount donated will be used to buy and delete CO2 quotas in the European quota-system, as well as donated to the UN Adaptation Fund. By buying and deleting CO2 quotas the emissions of CO2 is reduced. The researchers behind the survey will be in charge of these transactions. The amount used to buy CO2 quotas and donated to the UN Adaptation Fund is determined through your choice of climate policy, as well as the choices of the other participants. You can read more about CO2 quotas and the UN Adaptation Fund by following this link: https://www.adaptation-fund.org
>
> If you choose to donate, you have the option to receive certificates for the amount spent on buying and deleting CO2 quotas, and the donations to the UN Adaptation Fund, as documentation. For this purpose we will therefore ask for your email address later in the survey. It is your choice whether or not you wish to supply your email address or not. Remember that climate policy 1 and 2 always involve adaptation and also CO2 reduction if it is indicated in the description of the policy. Please also remember that the financing of the climate policy will be through a donation from you. Please make each of the 16 choices as if you had 200 DKK available each time.

As we can see, it may be rather lengthy to formulate if the policy context is intangible. Furthermore, it violates the general recommendation of not using donations as a payment vehicle as it may include other utility components such as warm-glow giving (Andreoni 1990). In conclusion, incentive compatibility is found to be important in empirical settings, yet DCE typically fails to satisfy its theoretical conditions due to the lack of single binary choices. It is always important to stress to respondents that choice sets are to be evaluated independently from each other. Furthermore, different ways to incentivise choices exist, e.g. with lotteries. While this is a possibility (see Palm-Forster et al. 2019 and Vossler and Zawojska 2020 for further discussion on the issue), it is not standard practice today.

2.5 Consequentiality

Consequentiality is defined by Carson and Groves (2007) as a situation in which a respondent thinks his/her answer can potentially influence the policy being investigated, whereby the answer to the survey is a possibility of influencing the policy—provided the policy is of interest to the respondent (see also Herriges et al. 2010 or Carson et al. 2014 for further discussion of this issue). This relates closely to the two first criteria mentioned with respect to incentive compatibility (see Sect. 2.4)—that respondents care about the outcome and that the payment can be enforced, but also to some extent with regard to the issue of binary questions—namely that it can be problematic to see an obvious outcome of several choices. The consequentiality issue in the literature is largely related to ensure that the answers to hypothetical questions can have an impact in the real world.

Within contingent valuation types of referenda, consequentiality has been investigated by, for example, Vossler and Evans (2009), who find that inconsequential questions lead to bias. The paper by Vossler et al. (2012) on consequentiality in DCE fundamentally shows that consequentiality on DCE is theoretically problematic as respondents answer multiple choice sets. Varying provision rules across split samples they find that this does not seem to be as important empirically.

Hassan et al. (2019) and Zawojska et al. (2019) distinguish between payment consequentiality and policy consequentiality, arguing that these two needs to be considered separately. Here, payment consequentiality is related to whether respondents believe that they will actually have to pay the cost of the chosen policy alternative if the policy is implemented in real life (i.e. free-riding is not possible). Policy consequentiality concerns whether respondents believe that their answers potentially influence the implementation of a policy, including whether the institution being paid has the institutional power to carry out the policy. In this regard, there is also the question of whether people trust that a policy has the described consequences, for example in terms of environmental improvements (Kassahun et al. 2016). Zawojska et al. (2019) find that policy and payment consequentiality have opposite effects on WTP, and therefore argue for them to be clearly distinguished and separately addressed.

For purely methodological DCE investigations, ensuring consequentiality may be challenging as the purpose of a study may be to learn more about the values of a certain good, but policy impact may be very far away. In these cases, it can be approached by telling respondents that the results of the survey will be communicated to politicians who may take it into account in their decision making. The more explicitly this can be done, the better. For example, it may be specified who will use this information and how it will be communicated. The more local a good is, and the more tangible, the easier it will often be to ensure such a communication and consideration.

If a survey is carried out on behalf of certain interest organisations or ministries (many studies are), policy consequentiality is often easier. However, highlighting the parties interested in the study may also lead to strategic answers (e.g. overbidding if an NGO is behind the survey with no power to force payments). So in such

cases specific awareness is to be given to payment consequentiality. Finally, respondents may distrust whether the stated environmental consequences will actually come about—i.e. outcome uncertainty. As we generally are after valuing an environmental improvement and not how this is obtained, it can lead to what we may call outcome or provision consequentiality (note, this is not a term used in the literature, and it is very similar in nature to the policy consequentiality described above). If such uncertainty is important—or important in people's minds, it will have to be addressed explicitly to avoid biasing the results by people's self-perceived probability estimates. Glenk and Colombo (2011) is an example in which an attribute is presented as uncertain and Lundhede et al. (2015) an example in which the policy is uncertain. One common approach to investigate perceptions of consequentiality is in follow-up questions, where we also test for strategic bidders and protest bidders (see Sects. 2.8 and 2.9). This means questions explicitly stating to what degree they think they would actually pay, to what degree they think politicians will be informed and take in the information (see, e.g., Oehlmann and Meyerhoff 2017).

In summary, consequentiality is mainly handled by the way the policy, the payment and the outcome are described. This has to be done in a clear way that is also perceived as realistic. Current practice is to highlight communication plans of the project in the survey to provide policy consequentiality. Payment consequentiality is described further in Sect. 2.10. Furthermore, follow-up questions on people's perception of consequentiality may be used. The importance of consequentiality and how to ensure it is still under development.

2.6 Cheap Talk, Opt-Out Reminder and Oath Script

One type of ex ante script that has received considerable attention in the literature is the so-called Cheap Talk script originally developed by Cummings and Taylor (1999) for use in a study based on a referendum Contingent Valuation Method (CVM). Cheap Talk explicitly describes the problem of hypothetical bias to respondents prior to the preference elicitation. In three independent contingent valuation surveys (Cummings and Taylor 1999) effectively eliminated hypothetical bias using a rather lengthy script of around 500 words which, firstly, described the hypothetical bias phenomena, secondly, outlined some possible explanations for it, and, finally, asked respondents to vote in the following hypothetical referendum as if it were real.

While these results initially suggested that using Cheap Talk would be an effective approach to avoid hypothetical bias, results from a wide range of subsequent studies testing Cheap Talk in various CVM settings are ambiguous (List 2001; Aadland and Caplan 2003, 2006; Lusk 2003; Murphy et al. 2005; Nayga et al. 2006; Champ et al. 2009; Morrison and Brown 2009; Barrage and Lee 2010; Mahieu 2010; Ladenburg et al. 2010; Ami et al. 2011; Carlsson et al. 2011). Similarly, empirical tests in DCE settings have found ambiguous effectiveness of Cheap Talk (List et al. 2006; Ozdemir et al. 2009; Carlsson et al. 2010; Silva et al. 2011; Tonsor and Shupp 2011; Bosworth and Taylor 2012; Moser et al. 2014; Howard et al. 2015). While there has

been no shortage of studies investigating Cheap Talk, it is relevant to note that most of these studies have used shorter scripts than the one originally used by Cummings and Taylor (1999). Despite the ambiguous results, it has become fairly common to include a Cheap Talk script when preparing questionnaires for empirical SP surveys. Exactly how common is difficult to assess, since details such as the inclusion or not of Cheap Talk (and other scripts) in questionnaires are not always reported when empirical survey results are published in scientific journals.

Johnston et al. (2017) note that the incentive properties of Cheap Talk are unclear, and it should thus not be applied without considering implications for framing and consequentiality. They further note that Cheap Talk directs the respondent's attention disproportionally to the costs, another aspect which requires caution. It would thus seem that Johnston et al. (2017) are generally sceptical towards using Cheap Talk in SP studies. However, it is not obvious that Cheap Talk as such is at odds with incentive compatibility and consequentiality. Considering the three overall parts of the full Cheap Talk script used by Cummings and Taylor (1999), the first part simply describing that people tend to overstate their WTP in hypothetical settings compared to real settings should not have adverse effects for incentive compatibility and consequentiality. The last part of the script, imploring respondents to answer as if it was a real choice situation, should also not have any adverse effects in this regard—on the contrary it encourages respondents to provide more truthful answers. The second part of the script, though, elaborating on possible reasons for hypothetical bias, could potentially be problematic if lack of incentive compatibility and/or lack of consequentiality are highlighted as potential reasons for hypothetical bias. As for the concern that Cheap Talk directs respondents' attention disproportionally towards the cost, it may be argued that this is exactly the purpose as hypothetical bias is essentially a result of respondents not paying as much attention to the cost in the hypothetical setting as they do in a non-hypothetical setting.

While few of the studies mentioned above have found Cheap Talk to completely remove hypothetical bias, most of them have found it to reduce hypothetical to at least some extent, and a few have found no effect at all. Only very few studies have found Cheap Talk to be outright counterproductive in terms of increasing hypothetical bias. Given that only two out of the more than 20 studies mentioned above find that using Cheap Talk actually leads to more biased WTP estimates than when Cheap Talk is not used, it would seem that for practical SP applications aimed at assessing WTP for non-marketed environmental goods the risk of introducing additional bias is outweighed by the greater chance of reducing bias. The actual impact will of course be context dependent and also depend on the specifics of the Cheap Talk script used. In relation to this it would seem that leaving out the second part of Cummings and Taylor's original script explaining the possible reasons for hypothetical bias would be favourable in order to avoid reducing survey consequentiality and incentive compatibility.

For self-administered survey modes, and in particular the increasingly used web surveys, where respondents due to limited attention budgets are likely to drop out or skip sections if faced with long text instructions (Lusk 2003; Bulte et al. 2005), using relatively short Cheap Talk scripts would seem preferable. These will typically be around 100 words in length. For example, a DCE-targeted short and neutral

Cheap Talk script which avoids explaining to respondents about possible reasons for hypothetical bias might read as follows:

> In surveys like this, we often find that some people tend to overestimate or underestimate how much they are actually willing to pay for implementation of alternative environmental policies. Thus, they may choose alternatives that they would not actually prefer in real life. It is important that your choices here are realistic. Hence, in each of the following choice tasks, please consider carefully that your household is actually able and willing to pay the costs associated with the alternative you choose.

Recognising first of all that Cheap Talk was originally developed for CVM, and secondly that CVM and DCE are inherently structurally different from each other, Ladenburg and Olsen (2014) proposed that Cheap Talk might not sufficiently address the specific structures of DCE that might be subject to hypothetical bias. One aspect where DCE differs structurally from CVM is that respondents commonly have to answer multiple choice tasks. Inspired by the fact that, for instance, anchoring effects in DCE have been shown to be transient over a sequence of choice tasks (Bateman et al. 2008; Ladenburg and Olsen 2008) and learning effects have also been shown to affect choice behaviour over a sequence of choice tasks (Carlsson et al. 2012). Ladenburg and Olsen (2014) speculate that the effect of Cheap Talk might be transient in DCE in the sense the effect would wear off after a few choice tasks since respondents would at some point forget about the reminder. Howard et al. (2015) confirm this suspicion. Ladenburg and Olsen (2014) thus suggest the use of a so-called Opt-Out Reminder.

The Opt-Out Reminder is a small script that explicitly reminds respondents to choose the opt-out alternative if they find the proposed experimentally designed alternatives in the choice set to be too expensive. An example of an Opt-Out Reminder for a DCE with a zero-priced opt-out alternative defined as a continuation of the current environmental policy is the following: "*If you find the environmental policy alternatives too expensive relative to the resulting improvements, you should choose the current policy*".

The Opt-Out Reminder is displayed just before each single choice set to account for the repeated choice nature of DCE. Ladenburg and Olsen (2014) found that adding the Opt-Out Reminder to a survey design which included Cheap Talk leads to significant reductions in WTP estimates. Varela et al. (2014) also tested the impact of presenting an Opt-Out Reminder together with Cheap Talk. Contrary to Ladenburg and Olsen (2014), the Opt-Out Reminder was not found to influence WTP. A possible explanation might be that Ladenburg and Olsen (2014) repeated the Opt-Out Reminder before each single choice set whereas Varela et al. (2014) only presented it once in the middle of the choice task sequence. This seems to support Ladenburg and Olsen (2014) who speculate that, given the repeated choice nature of DCE, it may be of particular importance to repeat the reminder since respondents might otherwise forget about the reminder as they progress through the choice tasks. A major limitation of both Ladenburg and Olsen (2014) and Varela et al. (2014) is that they test the Opt-Out Reminder in a purely hypothetical set-up. Thus, they cannot assess the degree of hypothetical bias mitigation since no fully incentivised treatment is conducted. In a recent study, Alemu and Olsen (2018) test the repeated Opt-Out

Reminder in an incentivised set-up where Cheap Talk is not included. They find that the Opt-Out Reminder effectively reduces hypothetical bias to a substantial degree, though not completely removing it for all attributes. More empirical tests of the reminder are obviously warranted before its general applicability can be thoroughly assessed.

While the incentive properties of the Opt-Out Reminder are not entirely clear, considering the fact that the simple and very short script essentially just reminds respondents to be rational at the extensive margin, it would not per se be at odds with incentive compatibility or consequentiality. Before applying the Opt-Out Reminder one should also consider whether it attracts disproportional attention to the cost attribute relative to other attributes. Again, seeing disproportional attention to non-cost attributes as a main driver of hypothetical bias, Ladenburg and Olsen (2014) developed the wording of the Opt-Out Reminder with the specific intention of drawing more attention to the cost attribute. It is not entirely obvious what disproportional attention refers to when mentioned in Johnston et al. (2017) in relation to Cheap Talk. A reasonable interpretation would seem to be that it is relative to attention in real or incentivised choice settings. Hence, the concern would be whether the Opt-Out Reminder makes respondents focus much more on the cost in the hypothetical choice experiment than they would in real life, essentially over-correcting for hypothetical bias. Assessing this of course requires real or incentivised choice settings with which to compare. So far, Alemu and Olsen (2018) is the only empirical study in this regard. They find that the Opt-Out Reminder does not over-correct for hypothetical bias, suggesting that it does not attract a disproportional amount of attention to the cost attribute.

Another more recently proposed ex ante approach that has shown some effect in terms of reducing hypothetical bias is the use of a so-called Oath Script or Honesty Priming exercises that encourage respondents to be truthful when stating their preferences. While the Oath Script directly asks respondents to swear an oath that they will truthfully answer the value eliciting questions, Honesty Priming is a somewhat more subtle approach that seeks to subconsciously prime respondents to answer truthfully but subjecting them to words that are associated with honesty. Carlsson et al. (2013), de-Magistris and Pascucci (2014), Jacquemet et al. (2013, 2017) and Stevens et al. (2013) find that the Oath Script effectively mitigates hypothetical bias. In a similar vein, de-Magistris et al. (2013) found Honesty Priming to mitigate hypothetical bias in a laboratory setting, but Howard et al. (2015) was not able to confirm this effect when testing this approach in a field setting.

The body of research investigating these approaches to induce honesty is far less extensive as is the case for Cheap Talk. Johnston et al. (2017) note that the behavioural impacts of these approaches are not yet well understood and may therefore have unintended consequences, and they basically end up recommending more research into this. This is underlined by the fact that these approaches are not (yet) commonly used in practice.

The NOAA panel (Arrow et al. 1993) strongly recommended reminding respondents both of relevant substitute commodities as well as budget constraints. They furthermore noted that this should be done forcefully and just before the valuation questions. In an empirical test, Loomis et al. (1994) found no impact of providing budget and substitute reminders. These findings, however, lead to a series of comments and replies in Land Economics in the years that followed (Whitehead and Blomquist 1999), indicating that there may be some effect from these reminders. Substantial literature has developed assessing the importance of substitute reminders, but mainly addressing it from a framing or embedding angle (Hailu et al. 2000; Rolfe et al. 2002; Jacobsen et al. 2011). It is not clear from the literature how much budget and substitute reminders have been used in practice, maybe because it has not been common practice to report the use of these reminders. Hailu et al. (2000) noted that few CVM studies had followed these NOAA recommendations up until the year 2000.

To sum up, there is no clear recommendation for DCE practitioners whether or not to use *ex ante* framing methods such as the above-mentioned Cheap Talk scripts, Opt-Out Reminders, Oath Scripts, Honesty Priming scripts, Budget Reminders or Substitute Reminders to reduce or eliminate hypothetical bias. For some of these, more investigations are needed in order to make solid conclusions, even though this is no guarantee for obtaining clear recommendations. Cheap Talk has been thoroughly scrutinised in the literature, but results are ambiguous, causing disagreement among DCE researchers concerning whether Cheap Talk should be used at all. At the end of the day, it is up to the DCE practitioner to decide on a case-by-case basis whether to use any of these *ex ante* framing methods. Ideally, if incentive compatibility and consequentiality has been ensured, hypothetical bias should not be a concern, and there would be no need for these approaches. However, in practice, in most cases it is not possible to secure these conditions in environmental DCE surveys, which means that hypothetical bias is likely to present a serious—and in most practical cases untestable—validity concern. In these cases, the practitioner should at least consider the pros and cons of the various *ex ante* framing methods, and for the particular empirical case and setting consider whether using one or more of them in combination is most likely to bring the elicited estimates of value closer to the true values or rather move them further away. Overall, the empirical evidence in the literature suggests that the latter rarely happens.

2.7 Instructional Choice Sets

Most people who respond to choice tasks in a DCE survey questionnaire are likely to face this kind of questionnaire for the first time. The unfamiliarity can mean that, at least among some respondents, the degree of randomness is larger in the first choice tasks than in subsequent ones (Carlsson et al. 2012). In a dichotomous choice CVM context, Carson et al. (1994) suggested providing respondents with "warm-up" choice tasks in order to reduce the experienced uncertainty related to unfamiliar

question context. Thus, in a couple of DCE surveys, respondents are presented a so-called instructional choice set (ICS) before they enter the sequence of choice tasks that will be used for estimating models and calculating WTP estimates (Ladenburg and Olsen 2008). Sometimes the former are also called training choice tasks while the latter are called value-elicitation choice tasks. The idea behind showing an ICS is to promote institutional learning, i.e. that respondents become more familiar with the choice context, the offered good and the choice tasks (see also Abate et al. 2018; Scheufele and Bennett 2012). The expected effect is that the ICS will reduce the degree of randomness that already exists among the first choices and thus improve the quality of choices recorded in the survey.

However, the literature has not provided clear evidence yet that the expected benefits from using an ICS, i.e. reducing the randomness of choices, will actually be achieved. At the same time, there are indications that the design of the ICS, especially the attribute level values shown on the ICS, might influence subsequent choices (e.g. Meyerhoff and Glenk 2015). This can happen, for example, because the attribute level values shown on the ICS might raise expectations and can, through anchoring, have an impact on all subsequent choices. Therefore, the overall effect of including an ICS might even be negative.

Given that the present evidence regarding the potential effects of an ICS is not yet conclusive, the following might be considered. Respondents generally get used to the choice task format quickly (Carlsson et al. 2012), an indicator of this is the often rapidly decreasing response time for a choice task as respondents move through the sequence of choice tasks (Meyerhoff and Glenk 2015). Thus, an ICS might not be that important. Instead, it is important that the order of appearance of the choice tasks a respondent faces is randomised. This way, not all respondents will have the same choice task as their first task and the potential anchoring effects, while still present, will now be dispersed over the full range of attribute levels rather than attached to one specific set of attribute levels. Even if the design is blocked, i.e. those who are assigned to different blocks face a different first choice task, it is essential to randomise the order of appearance to even up potential ordering effects.

If, however, an ICS is considered indispensable for reducing, for example, institutional uncertainty as the choice tasks are very complex and difficult to capture, the attribute level values shown on the ICS should be selected carefully. As these values could affect subsequent choices it seems advisable to avoid extreme values, i.e. the attribute level values representing the worst or best quality in your design, and level values in the middle of your attribute level range should be used. Especially very low values for cost in combination with levels of non-monetary attributes representing high quality levels might raise expectations that "you don't have to spend a lot in order to get high quality". Therefore, respondents might wait for similar alternatives and, as a consequence, the share of status quo choices could increase. In the event the attribute displays one of the highest levels available from among the range of cost levels, it could mean that those who have seen high price levels on the ICS do not trust that good quality can also be achieved at low costs. One option is to randomly select the attribute levels on the ICS so that numerous respondents see differently composed ICS. Another way to mitigate potential effects of anchoring could be to

randomly draw a choice task from your experimental design so that the ICS differs across all respondents (Uggeldahl 2018).

2.8 Identifying Protesters

Protest responses are those systematically choosing the status quo option in a DCE, thus rejecting or protesting against some aspect of the constructed market scenario (Meyerhoff and Liebe 2006). In order to detect them, follow-up questions on the reasons for their answers are usually added to the valuation questions. Given that protest responses may lead to inconsistent welfare estimation, the researcher should properly detect and treat them (Meyerhoff and Liebe 2010).

Common identification approaches to protest answers include the use of debriefing questions, statistical outlier analysis and identification of systematic patterns in a set of choice situations. A typical debriefing question to identify zero protesters is to present a list of predetermined statements to respondents who consistently choose the zero-priced opt-out alternative throughout the choice set sequence, and ask which of these statements best corresponds to the reason why they always chose the opt-out. The list should include a range of statements of which some should indicate valid zeros (i.e. the choices are made in line with random utility maximisation, reflecting true preferences) while others should indicate protest zero (i.e. the choices made do not reflect the respondent's true preferences for the described good). When developing the list of statements, it is important to carefully consider the interpretation and classification of each statement to avoid ambiguous statements that afterwards cannot be clearly classified as a protest or valid. Usually protest answers take the form of beliefs that others (governments, private companies, etc.) are responsible and that they should bear the costs. As an example, the following statements can be used; ultimately they may depend on the context (Table 2.2).

Some authors have suggested it is better to use open-ended questions for the motives behind protest answers and then code them, which could lower the protest rate

Table 2.2 Items used for identifying protesters

Indicating protest	Indicating valid zero WTP
I do not want to put a dollar value on protecting nature	Society has more important problems than protecting plants and animals
Someone else should pay to protect natural resources	I cannot afford to pay
Not enough information is given	I do not think protecting plants and animals is worth the specified price
I object to the way the question is asked	I already spend a lot on animal welfare initiatives
The payment method is inappropriate	

Adapted from Meyerhoff and Liebe (2006, 2010) and Frey and Pirscher (2019)

(see, e.g., Bateman et al. 2002). It is also important to distinguish protest responses (e.g. "I don't want to put a dollar value on protecting plants and animals" from genuine zeros (e.g. "I can't afford to pay" or "I don't want to pay"). It has also been argued that those who are willing to pay may still hold some protest beliefs (Jorgensen and Syme 2000; Meyerhoff and Liebe 2006).

Statistical treatment of these responses includes dropping observations, sensitivity analysis to determine their impact on welfare measures, or using specific choice models able to accommodate protest responses (e.g. Meyerhoff et al. 2012). Sample selection models have also been proposed in order to take into account both zero values and protest answers in the model estimates (Strazzera et al. 2003; Grammatikopoulou and Olsen 2013). Glenk et al. (2012) argue that the latent class approach to modelling non-participation requires an absence of a priori assumptions about how to "treat" protest responses and serial non-participation, and has the advantage over alternative approaches such as double hurdle choice models (e.g. von Haefen et al. 2005) that it does not require a priori identification of non-participation (Burton and Rigby 2009). It is important to denote that identifying protest responses does not necessarily imply a binary (yes/no) treatment. Practitioners should always keep in mind that the way protesters are handled could significantly influence welfare measures.

Whether protest responses should be included or excluded from the data analysis remains an open question. While many applications tend to exclude them, others have argued that, in order to provide more conservative estimates of WTP, protest answers should be included in the data analysis (Carson and Hanemann 2005). As there is no agreement on the best treatment for protest responses, transparency both in detection and treatment of these responses is found to be essential (Johnston et al. 2017), especially if the DCE is conducted for policy purposes.

More specifically, practitioners should collect different reasons for opting-out or systematically choosing the status quo option, including both protest and other reasons, and comprehensively report on this matter, including the overall number of protesters (frequency and percentages), the method employed to determine them (open-ended versus attitudinal questions) and the influence on welfare estimates of including/excluding protest responses (Meyerhoff and Liebe 2010).

2.9 Identifying Strategic Bidders

Strategic behaviour occurs when respondents do not answer truthfully to the valuation questions of a survey because they think that they can affect the final outcome of the survey by answering differently (Hoyos and Mariel 2010). For example, they could answer affirmatively to a high price, showing that the good is very valuable although thinking that they will never have to pay it in reality. As seen in Sect. 2.4, this could be the case when the survey lacks incentive compatibility and payment consequentiality.

Strategic behaviour from respondents has been used as a general criticism to mistrust SP survey responses, especially in the contingent valuation literature (Carson and Hanemann 2005; McFadden and Train 2017). However, strategic bias in empirical studies can be minimised through well-designed questionnaires (Mitchell and Carson 1989). For example, consistent with a lack of incentive compatibility and potentially inducing strategic behaviour, the use of open-ended questions in contingent valuation has decreased in recent years relative to other formats, in part due to the large number of respondents who provide either unrealistically high or zero WTP responses. Respondents with apparently extreme sensitivities can also be accommodated in discrete choice models (Campbell et al. 2010).

DCE have been found to help avoid strategic behaviour from the respondents (Hanley et al. 2001; Lancsar and Louviere 2008), although some authors like Day et al. (2012) find empirical evidence of strategic behaviour in the context of valuing public goods. Nonetheless, the researcher should bear in mind that the advanced disclosure of choice tasks often involved in a typical DCE as well as presenting multiple choice tasks and alternatives or even the order in which they are presented (i.e. departing from the theoretically incentive compatible single binary choice between the status quo and one alternative) could induce strategic behaviour from respondents (Collins and Vossler 2009; McNair et al. 2011; Vossler et al. 2012; Scheufele and Bennett 2012).

From the previous discussion, it is clearly unlikely that strategic responses will be avoided, so practitioners should try to minimise them by using incentive compatible choice experiments and plausible consequential decision setting while considering the use of other methods to minimise hypothetical bias.

2.10 Payment Vehicle and Cost Vector Design

The payment vehicle used can be anything from which respondents experience a negative utility (in a WTP setting, or a positive in a WTA setting). The crucial point is that it has to be considered realistic, relevant and consequential by the respondent. Thus, it relies heavily on the institutional context of a given country. In choosing the right payment vehicle, it is important to ensure a mandatory payment if used (see Sect. 2.5 on consequentiality), that it is a vehicle that is available to the respondents, and that the vehicle match the type of good. For example, if we are dealing with public good aspects of water, a water consumption user fee may lead to a high level of protesters, even if it is the only realistically available and mandatory payment vehicle (see Sect. 2.8). Hassan et al. (2018) have a thorough discussion of the choice of payment vehicle in a case where the choice was not so obvious.

The most common payment vehicles typically involve some kind of monetary transfer. Examples of payment vehicles in a utility enhancing context include income tax (Campbell et al. 2014), tax on water usage (Jørgensen et al. 2013), subsidy reduction (Hassan et al. 2018), entrance fee (Talpur et al. 2018), and in a utility decreasing context, subsidies paid to landowners (Vedel et al. 2015a), donations

from NGOs (Rakotonarivo et al. 2017), lowering property tax (Vedel et al. 2015b), salaries from alternative employment (Nielsen et al. 2014), the opportunity gain of an interest free loan or labour (Kassahun and Jacobsen 2015). The choice of payment vehicle should always be guided and thoroughly tested in focus group interviews. In particular, it is important to ascertain that people consider the chosen payment vehicle to be both realistic, relevant and consequential for the specific valuation context.

Once the payment vehicle is decided on, an appropriate cost vector has to be determined. The lower bound of the cost vector will typically be logically located at zero if the survey aims at willingness-to-pay estimates. In the case of willingness to accept (WTA) estimates, of course, the cost levels might be negative indicating a discount, for example. Referring back to the dichotomous choice contingent valuation literature, greater effort is required to identify the upper end of the range. This is the so-called choke price, i.e. a payment level that is so high that it just chokes off (almost) any demand for the offered improvement—essentially the price at which the demand curve reaches zero. In the DCE context, this corresponds to a payment level at which almost no one (a commonly used rule of thumb is that it should be less than 5%) would choose the presented alternative regardless of the other attribute levels of the alternative and other available alternatives in the choice set. Once this upper bound of the cost vector has been found, a suitable number and location of levels needs to be set within the chosen lower and upper bound for the cost vector. Sufficiently high cost levels are particularly important for identifying respondents with a very low cost sensitivity who are typically situated in the tail of the distribution.

Concerning the number of levels of the cost attribute, narrowly focusing on D-efficiency is likely to lead to relatively few levels. Nevertheless, this may not be optimal given the importance of the estimated cost attribute parameter for the calculation of all WTP estimates. From this point of view, the cost parameter should be estimated with the highest possible precision, also for smaller level changes and to allow for possible nonlinear preferences. No fixed number of levels can be a priori recommended, but most practical applications of environmental DCE use more levels for the cost attribute than for non-cost attributes. Typically between 4 and 8 levels in addition to zero are used for the cost attribute. Next, the location of the levels within the range also needs to be determined. This could be done by distributing the levels evenly within the range. An example is the following cost vector with seven levels: {0; 10; 20; 30; 40; 50; 60}. Rather than using such equidistant spacing of levels in the cost vector, a more commonly used approach is to use (approximately) exponentially increasing distance between levels. An example using the same range as above is: {0; 2; 4; 8; 15; 30; 60}. However, to our knowledge no systematic investigation of the pros and cons of both approaches is available, as is also the problem for other aspects of the cost vector design mentioned above. When linear utility functions are used it may be beneficial to implement unequally spaced cost levels to increase the number of cost differences across the attributes, thereby facilitating the estimation of the cost coefficient, its heterogeneity and WTP measures accordingly.

Another decision to take is whether there should be non-status-quo alternatives with a price of zero. From a statistical point of view this may be wise, especially if the sign of some of the attributes may differ, but it may be problematic to ensure

policy consequentiality if improvements can be obtained at no cost. Both approaches are found in the literature. In theory, the levels and range of the cost vector should not matter. The problem in a DCE context is, however, that there is evidence that people may anchor their choices in the payment levels and range presented (Glenk et al. 2019). For instance, Kragt (2013) analyses the importance of the bid range by a split sample where one split got a bid range from AU$ 0–400 and the other from AU$ 0–600, both with 5 levels in each. She concludes that respondents anchor their choices to relative bid levels; yet she finds little effect on the actual WTP. Other similar studies (e.g. Hanley et al. 2005; Carlsson and Martinsson 2008) find ambiguous evidence regarding impacts on WTP. More recently, Glenk et al. (2019) find WTP estimates to be significantly affected by the payment vector. Furthermore, Mørkbak et al. (2010) find that the specific choke price used may affect WTP estimates.

In conclusion, choosing the right payment vehicle is important to ensure consequentiality and thus validity of the study. It has to be broadly accepted in the population, mandatory for all to contribute to and there has to be trust in the institutional setting. This is typically identified and tested in focus groups, and further validated by follow-up questions after the choice sets in the survey. Despite the importance of the cost vector, there are few solid recommendations for determining an appropriate cost vector in practice—partly because it is highly context dependent. Identification of the cost vector should thus always be guided by inputs from focus group interviews. Furthermore, it should be ascertained in pilot tests that the cost attribute parameter can be estimated with a high level of statistical significance and that alternatives displaying the highest level of the cost vector are only very rarely chosen.

References

Aadland D, Caplan AJ (2003) Willingness to pay for curbside recycling with detection and mitigation of hypothetical bias. Am J Agr Econ 85:492–502. https://doi.org/10.1111/1467-8276.00136

Aadland D, Caplan AJ (2006) Cheap talk reconsidered: new evidence from CVM. J Econ Behav Organ 60:562–578. https://doi.org/10.1016/j.jebo.2004.09.006

Aanesen M, Armstrong C, Czajkowski M et al (2015) Willingness to pay for unfamiliar public goods: preserving cold-water coral in Norway. Ecol Econ 112:53–67. https://doi.org/10.1016/j.ecolecon.2015.02.007

Abate TG, Mørkbak MR, Olsen SB (2018) Inducing value and institutional learning effects in stated choice experiments using advanced disclosure and instructional choice set treatments. Agric Econ 49:339–351. https://doi.org/10.1111/agec.12420

Alemu MH, Olsen SB (2018) Can a repeated opt-out reminder mitigate hypothetical bias in discrete choice experiments? An application to consumer valuation of novel food products. Eur Rev Agric Econ 45:749–782. https://doi.org/10.1093/erae/jby009

Ami D, Aprahamian F, Chanel O, Luchini S (2011) A Test of cheap talk in different hypothetical contexts: the case of air pollution. Environ Resource Econ 50:111. https://doi.org/10.1007/s10640-011-9464-z

Andreoni J (1990) impure altruism and donations to public goods: a theory of warm-glow giving. Econ J 100:464–477. https://doi.org/10.2307/2234133

Arrow K, Solow R, Portney P et al (1993) Report of NOAA Panel on contingent valuation. Fed Reg 58:4601–4614

Barrage L, Lee MS (2010) A penny for your thoughts: inducing truth-telling in stated preference elicitation. Econ Lett 106:140–142. https://doi.org/10.1016/j.econlet.2009.11.006

Bateman IJ, Burgess D, Hutchinson WG, Matthews DI (2008) Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness. J Environ Econ Manag 55:127–141. https://doi.org/10.1016/j.jeem.2007.08.003

Bateman IJ, Carson RT, Day BH et al (2002) Economic valuation with stated preferences techniques: a manual. Edward Elgar, Cheltenham

Bateman IJ, Day BH, Jones AP, Jude S (2009) Reducing gain–loss asymmetry: a virtual reality choice experiment valuing land use change. J Environ Econ Manag 58:106–118. https://doi.org/10.1016/j.jeem.2008.05.003

Bishop RC, Boyle KJ, Carson RT et al (2017) Putting a value on injuries to natural assets: the BP oil spill. Science 356:253–254. https://doi.org/10.1126/science.aam8124

Bosworth R, Taylor LO (2012) Hypothetical bias in choice experiments: is cheap talk effective at eliminating bias on the intensive and extensive margins of choice? BE J Econ Anal Policy 12. https://doi.org/10.1515/1935-1682.3278

Bulte E, Gerking S, List JA, de Zeeuw A (2005) The effect of varying the causes of environmental problems on stated WTP values: evidence from a field study. J Environ Econ Manag 49:330–342. https://doi.org/10.1016/j.jeem.2004.06.001

Burton M, Rigby D (2009) Hurdle and latent class approaches to serial non-participation in choice models. Environ Resource Econ 42:211–226. https://doi.org/10.1007/s10640-008-9225-9

Campbell D, Hensher DA, Scarpa R (2014) Bounding WTP distributions to reflect the "actual" consideration set. J Choice Model 11:4–15. https://doi.org/10.1016/j.jocm.2014.02.004

Campbell D, Hess S, Scarpa R, Rose JM (2010) Accommodating coefficient outliers in discrete choice modelling: a comparison of discrete and continuous mixing approaches. In: Hess S, Daly A (eds) Choice modelling: The State-of-the-art and The State-of-practice. Emerald Group Publishing Limited, pp 331–352

Carlsson F, García JH, Löfgren Å (2010) Conformity and the demand for environmental goods. Environ Resource Econ 47:407–421. https://doi.org/10.1007/s10640-010-9385-2

Carlsson F, Kataria M, Krupnick A et al (2013) The truth, the whole truth, and nothing but the truth—a multiple country test of an oath script. J Econ Behav Organ 89:105–121. https://doi.org/10.1016/j.jebo.2013.02.003

Carlsson F, Martinsson P (2008) Does it matter when a power outage occurs?—a choice experiment study on the willingness to pay to avoid power outages. Energy Econ 30:1232–1245. https://doi.org/10.1016/j.eneco.2007.04.001

Carlsson F, Martinsson P, Akay A (2011) The effect of power outages and cheap talk on willingness to pay to reduce outages. Energy Econ 33:790–798. https://doi.org/10.1016/j.eneco.2011.01.004

Carlsson F, Mørkbak MR, Olsen SB (2012) The first time is the hardest: a test of ordering effects in choice experiments. J Choice Model 5:19–37. https://doi.org/10.1016/S1755-5345(13)70051-4

Carson RT (2000) Contingent valuation: a user's guide. Environ Sci Technol 34:1413–1418. https://doi.org/10.1021/es990728j

Carson RT, Groves T (2007) Incentive and informational properties of preference questions. Environ Resource Econ 37:181–210. https://doi.org/10.1007/s10640-007-9124-5

Carson RT, Groves T, List JA (2014) Consequentiality: a theoretical and experimental exploration of a single binary choice. J Assoc Environ Resource Econ 1:171–207. https://doi.org/10.1086/676450

Carson RT, Hanemann W (2005) Chapter 17 Contingent Valuation. Handbook of Environmental Economics 2:821–936. https://doi.org/10.1016/S1574-0099(05)02017-6

Champ PA, Moore R, Bishop RC (2009) A Comparison of approaches to mitigate hypothetical bias. Agric Resource Econ Rev 38:166–180. https://doi.org/10.1017/S106828050000318X

Carson RT, Louviere JJ, Anderson DA et al (1994) Experimental analysis of choice. Market Lett 5:351–367. https://doi.org/10.1007/BF00999210

Chilton SM, Hutchinson WG (1999) Do focus groups contribute anything to the contingent valuation process? J Econ Psychol 19

Collins JP, Vossler CA (2009) Incentive compatibility tests of choice experiment value elicitation questions. J Environ Econ Manag 58:226–235. https://doi.org/10.1016/j.jeem.2009.04.004

Cummings RG, Taylor LO (1999) Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. Am Econ Rev 89:649–665. https://doi.org/10.1257/aer.89.3.649

Day B, Bateman IJ, Carson RT et al (2012) Ordering effects and choice set awareness in repeat-response stated preference studies. J Environ Econ Manag 63:73–91. https://doi.org/10.1016/j.jeem.2011.09.001

de Magistris T, Gracia A, Nayga RM (2013) On the use of honesty priming tasks to mitigate hypothetical bias in choice experiments. Am J Agric Econ 95:1136–1154. https://doi.org/10.1093/ajae/aat052

de-Magistris T, Pascucci S (2014) The effect of the solemn oath script in hypothetical choice experiment survey: a pilot study. Economics Letters 123:252–255. https://doi.org/10.1016/j.econlet.2014.02.016

Dillman DA, Smyth JD, Christian LM (2008) Internet, mail, and mixed-mode Surveys: the tailored design method, 3rd Revised edn. Wiley, Hoboken, NJ

Dillman DA, Smyth JD, Christian LM (2014) Internet, mail, and mixed-mode surveys: the tailored design method, 4th edn. Wiley, Hoboken, NJ

Duncan GJ, Petersen E (2001) The long and short of asking questions about income, wealth, and labor supply. Soc Sci Res 30:248–263. https://doi.org/10.1006/ssre.2000.0696

Frey UJ, Pirscher F (2019) Distinguishing protest responses in contingent valuation: a conceptualization of motivations and attitudes behind them. PLoS ONE 14:e0209872. https://doi.org/10.1371/journal.pone.0209872

Galesic M, Bosnjak M (2009) Effects of questionnaire length on participation and indicators of response quality in a web survey. Public Opin Q 73:349–360. https://doi.org/10.1093/poq/nfp031

Glenk K, Colombo S (2011) How sure can you be? A framework for considering delivery uncertainty in benefit assessments based on stated preference methods. J Agric Econ 62:25–46. https://doi.org/10.1111/j.1477-9552.2010.00278.x

Glenk K, Hall C, Liebe U, Meyerhoff J (2012) Preferences of Scotch malt whisky consumers for changes in pesticide use and origin of barley. Food Policy 37:719–731. https://doi.org/10.1016/j.foodpol.2012.08.003

Glenk K, Meyerhoff J, Akaichi F, Martin-Ortega J (2019) Revisiting cost vector effects in discrete choice experiments. Resource Energy Econ 57:135–155. https://doi.org/10.1016/j.reseneeco.2019.05.001

Grammatikopoulou I, Olsen SB (2013) Accounting protesting and warm glow bidding in Contingent Valuation surveys considering the management of environmental goods—an empirical case study assessing the value of protecting a Natura 2000 wetland area in Greece. J Environ Manage 130:232–241. https://doi.org/10.1016/j.jenvman.2013.08.054

Groves RM, Presser S, Dipko S (2004) the role of topic interest in survey participation decisions. Public Opin Q 68:2–31. https://doi.org/10.1093/poq/nfh002

Hailu A, Adamowicz WL, Boxall PC (2000) Complements, substitutes, budget constraints and valuation: application of a multi-program environmental valuation method. Environ Resource Econ 16:51–68. https://doi.org/10.1023/A:1008328920083

Hanley N, Adamowicz W, Wright RE (2005) Price vector effects in choice experiments: an empirical test. Resource Energy Econ 27:227–234. https://doi.org/10.1016/j.reseneeco.2004.11.001

Hanley N, Mourato S, Wright RE (2001) Choice modelling approaches: a superior alternative for environmental valuation? J Econ Surv 15:435–462. https://doi.org/10.1111/1467-6419.00145

Hassan S, Olsen SB, Thorsen BJ (2019) Urban-rural divides in preferences for wetland conservation in Malaysia. Land Use Policy 84:226–237. https://doi.org/10.1016/j.landusepol.2019.03.015

Hassan S, Olsen SB, Thorsen BJ (2018) Appropriate payment vehicles in stated preference studies in developing economies. Environ Resource Econ 71:1053–1075. https://doi.org/10.1007/s10640-017-0196-6

Herriges J, Kling C, Liu C-C, Tobias J (2010) What are the consequences of consequentiality? J Environ Econ Manag 59:67–81. https://doi.org/10.1016/j.jeem.2009.03.004

Howard G, Roe BE, Nisbet EC, Martin J (2015) Hypothetical bias mitigation in choice experiments: effectiveness of cheap talk and honesty priming fade with repeated choices. Social Science Research Network, Rochester, NY

Hoyos D (2010) The state of the art of environmental valuation with discrete choice experiments. Ecol Econ 69:1595–1603. https://doi.org/10.1016/j.ecolecon.2010.04.011

Hoyos D, Mariel P (2010) Contingent valuation: past, present and future. Prague Economic Papers, 329–343. https://doi.org/10.18267/j.pep.380

Jacobsen JB, Boiesen JH, Thorsen BJ, Strange N (2008) What's in a name? The use of quantitative measures versus 'Iconised' species when valuing biodiversity. Environ Resource Econ 39:247–263. https://doi.org/10.1007/s10640-007-9107-6

Jacobsen JB, Lundhede TH, Martinsen L et al (2011) Embedding effects in choice experiment valuations of environmental preservation projects. Ecol Econ 70:1170–1177. https://doi.org/10.1016/j.ecolecon.2011.01.013

Jacquemet N, James A, Luchini S, Shogren JF (2017) Referenda under oath. Environ Resource Econ 67:479–504. https://doi.org/10.1007/s10640-016-0023-5

Jacquemet N, Joule R-V, Luchini S, Shogren JF (2013) Preference elicitation under oath. J Environ Econ Manag 65:110–132. https://doi.org/10.1016/j.jeem.2012.05.004

Jensen AK, Johnston RJ, Olsen SB (2019) Does one size really fit all? Ecological endpoint heterogeneity in stated preference welfare analysis. Land Econ 95:307–332. https://doi.org/10.3368/le.95.3.307

Johnston RJ, Boyle KJ, Adamowicz W (Vic) et al (2017) Contemporary guidance for stated preference studies. J Assoc Environ Resource Econ 4:319–405. https://doi.org/10.1086/691697

Jorgensen BS, Syme GJ (2000) Protest responses and willingness to pay: attitude toward paying for stormwater pollution abatement. Ecol Econ 33:251–265. https://doi.org/10.1016/S0921-8009(99)00145-7

Jørgensen SL, Olsen SB, Ladenburg J et al (2013) Spatially induced disparities in users' and non-users' WTP for water quality improvements—testing the effect of multiple substitutes and distance decay. Ecol Econ 92:58–66. https://doi.org/10.1016/j.ecolecon.2012.07.015

Kaplowitz MD, Hoehn JP (2001) Do focus groups and individual interviews reveal the same information for natural resource valuation? Ecol Econ 36:237–247. https://doi.org/10.1016/S0921-8009(00)00226-3

Kassahun HT, Jacobsen JB (2015) Economic and institutional incentives for managing the Ethiopian highlands of the Upper Blue Nile Basin: a latent class analysis. Land Use Policy 44:76–89. https://doi.org/10.1016/j.landusepol.2014.11.017

Kassahun HT, Nicholson CF, Jacobsen JB, Steenhuis TS (2016) Accounting for user expectations in the valuation of reliable irrigation water access in the Ethiopian highlands. Agric Water Manag 168:45–55. https://doi.org/10.1016/j.agwat.2016.01.017

Kragt ME (2013) The effects of changing cost vectors on choices and scale heterogeneity. Environ Resource Econ 54:201–221. https://doi.org/10.1007/s10640-012-9587-x

Krueger RA, Casey MA (2008) Focus Groups: a practical guide for applied research, 4th edn. Sage, Los Angeles

Ladenburg J, Dahlgaard JO, Bonnichsen O (2010) Testing the effect of a short cheap talk script in choice experiments. University of Copenhagen, Department of Food and Resource Economics

Ladenburg J, Olsen SB (2014) Augmenting short cheap talk scripts with a repeated opt-out reminder in choice experiment surveys. Resource Energy Econ 37:39–63. https://doi.org/10.1016/j.reseneeco.2014.05.002

Ladenburg J, Olsen SB (2008) Gender-specific starting point bias in choice experiments: evidence from an empirical study. J Environ Econ Manag 56:275–285. https://doi.org/10.1016/j.jeem.2008.01.004

Lancsar E, Louviere J (2008) Conducting discrete choice experiments to inform healthcare decision making: a user's guide. PharmacoEconomics 26:661–677. https://doi.org/10.2165/00019053-200826080-00004

Leeuw EDD, Hox J, Dillman D (2008) International handbook of survey methodology, 1st edn. Routledge, New York, London

Liebe U, Hundeshagen C, Beyer H, von Cramon-Taubadel S (2016) Context effects and the temporal stability of stated preferences. Soc Sci Res 60:135–147. https://doi.org/10.1016/j.ssresearch.2016.04.013

Lim SL, Yang J-C, Ehrisman J et al (2020) Are videos or text better for describing attributes in stated-preference surveys? Patient. https://doi.org/10.1007/s40271-020-00416-9

List JA (2001) Do explicit warnings eliminate the hypothetical bias in elicitation procedures? evidence from field auctions for sportscards. Am Econ Rev 91:1498–1507. https://doi.org/10.1257/aer.91.5.1498

List JA, Sinha P, Taylor MH (2006) Using choice experiments to value non-market goods and services: evidence from field experiments. BE J Econ Anal Policy 6. https://doi.org/10.2202/1538-0637.1132

Loomis J, Gonzalez-Caban A, Gregory R (1994) Do reminders of substitutes and budget constraints influence contingent valuation estimates? Land Econ 70:499–506. https://doi.org/10.2307/3146643

Lundhede T, Jacobsen JB, Hanley N et al (2015) Incorporating outcome uncertainty and prior outcome beliefs in stated preferences. Land Economics 91:296–316. https://doi.org/10.3368/le.91.2.296

Lusk JL (2003) Effects of cheap talk on consumer willingness-to-pay for golden rice. Am J Agric Econ 85:840–856. https://doi.org/10.1111/1467-8276.00492

Mahieu P-A (2010) Does gender matter when using cheap talk in contingent valuation studies? Econ Bull 30:2955–2961

Matthews Y, Scarpa R, Marsh D (2017) Using virtual environments to improve the realism of choice experiments: A case study about coastal erosion management. J Environ Econ Manag 81:193–208. https://doi.org/10.1016/j.jeem.2016.08.001

McFadden D, Train K (2017) Contingent valuation of environmental goods: a comprehensive critique. Edward Elgar Publishing

McNair BJ, Bennett J, Hensher DA, Rose JM (2011) Households' willingness to pay for overhead-to-underground conversion of electricity distribution networks. Energy Policy 39:2560–2567. https://doi.org/10.1016/j.enpol.2011.02.023

Meyerhoff J, Bartczak A, Liebe U (2012) Protester or non-protester: a binary state? On the use (and non-use) of latent class models to analyse protesting in economic valuation. Aust J Agric Resource Econ 56:438–454. https://doi.org/10.1111/j.1467-8489.2012.00582.x

Meyerhoff J, Glenk K (2015) Learning how to choose-effects of instructional choice sets in discrete choice experiments. Resource Energy Econ 41:122–142. https://doi.org/10.1016/j.reseneeco.2015.04.006

Meyerhoff J, Liebe U (2006) Protest beliefs in contingent valuation: explaining their motivation. Ecol Econ 57:583–594. https://doi.org/10.1016/j.ecolecon.2005.04.021

Meyerhoff J, Liebe U (2010) Determinants of protest responses in environmental valuation: a meta-study. Ecol Econ 70:366–374. https://doi.org/10.1016/j.ecolecon.2010.09.008

Mitchell RC, Carson RT (1989) Using surveys to value public goods: the contingent valuation method. RFF Press, Washington, D.C.

Moore DW (2002) Measuring new types of question-order effects: additive and subtractive. Public Opin Q 66:80–91. https://doi.org/10.1086/338631

Morgan DL (1997) Focus groups as qualitative research, 1st edn. Sage, Thousand Oaks, CA

Mørkbak MR, Christensen T, Gyrd-Hansen D (2010) Choke price bias in choice experiments. Environ Resource Econ 45:537–551. https://doi.org/10.1007/s10640-009-9327-z

Morrison M, Brown TC (2009) Testing the effectiveness of certainty scales, cheap talk, and dissonance-minimization in reducing hypothetical bias in contingent valuation studies. Environ Resource Econ 44:307–326. https://doi.org/10.1007/s10640-009-9287-3

Moser R, Raffaelli R, Notaro S (2014) Testing hypothetical bias with a real choice experiment using respondents' own money. Eur Rev Agric Econ 41:25–46. https://doi.org/10.1093/erae/jbt016

Murphy JJ, Allen PG, Stevens TH, Weatherhead D (2005) A meta-analysis of hypothetical bias in stated preference valuation. Environ Resource Econ 30:313–325

Nayga RM, Woodward R, Aiew W (2006) Willingness to pay for reduced risk of foodborne illness: a nonhypothetical field experiment. Can J Agric Econ/Revue Canadienne D'agroeconomie 54:461–475. https://doi.org/10.1111/j.1744-7976.2006.00061.x

Nielsen ASE, Lundhede TH, Jacobsen JB (2016) Local consequences of national policies—a spatial analysis of preferences for forest access reduction. Forest Policy Econ 73:68–77. https://doi.org/10.1016/j.forpol.2016.08.010

Nielsen MR, Jacobsen JB, Thorsen BJ (2014) Factors determining the choice of hunting and trading bushmeat in the Kilombero Valley, Tanzania. Conserv Biol 28:382–391. https://doi.org/10.1111/cobi.12197

Oehlmann M, Meyerhoff J (2017) Stated preferences towards renewable energy alternatives in Germany—do the consequentiality of the survey and trust in institutions matter? J Environ Econ Policy 6:1–16. https://doi.org/10.1080/21606544.2016.1139468

Ozdemir S, Johnson FR, Hauber AB (2009) Hypothetical bias, cheap talk, and stated willingness to pay for health care. J Health Econ 28:894–901. https://doi.org/10.1016/j.jhealeco.2009.04.004

Palm-Forster LH, Ferraro PJ, Janusch N et al (2019) Behavioral and experimental agri-environmental research: methodological challenges, literature gaps, and recommendations. Environ Resource Econ 73:719–742. https://doi.org/10.1007/s10640-019-00342-x

Patterson Z, Darbani JM, Rezaei A et al (2017) Comparing text-only and virtual reality discrete choice experiments of neighbourhood choice. Landscape Urban Plan 157:63–74. https://doi.org/10.1016/j.landurbplan.2016.05.024

Pouta E (2004) Attitude and belief questions as a source of context effect in a contingent valuation survey. J Econ Psychol 25:229–242. https://doi.org/10.1016/S0167-4870(02)00170-8

Rakotonarivo OS, Jacobsen JB, Larsen HO et al (2017) Qualitative and quantitative evidence on the true local welfare costs of forest conservation in Madagascar: are discrete choice experiments a valid ex ante tool? World Dev 94:478–491. https://doi.org/10.1016/j.worlddev.2017.02.009

Rid W, Haider W, Ryffel A, Beardmore B (2018) Visualisations in choice experiments: comparing 3D film-sequences and still-images to analyse housing development alternatives. Ecol Econ 146:203–217. https://doi.org/10.1016/j.ecolecon.2017.10.019

Rolfe J, Bennett J, Louviere J (2002) Stated values and reminders of substitute goods: testing for framing effects with choice modelling. Aust J Agric Resource Econ 46:1–20. https://doi.org/10.1111/1467-8489.00164

Rossetti T, Hurtubia R (2020) An assessment of the ecological validity of immersive videos in stated preference surveys. J Choice Model 34:100198. https://doi.org/10.1016/j.jocm.2019.100198

Sandorf ED, Aanesen M, Navrud S (2016) Valuing unfamiliar and complex environmental goods: a comparison of valuation workshops and internet panel surveys with videos. Ecol Econ 129:50–61. https://doi.org/10.1016/j.ecolecon.2016.06.008

Scheufele G, Bennett J (2012) Response strategies and learning in discrete choice experiments. Environ Resource Econ 52:435–453. https://doi.org/10.1007/s10640-011-9537-z

Schkade DA, Payne JW (1994) How people respond to contingent valuation questions: a verbal protocol analysis of willingness to pay for an environmental regulation. J Environ Econ Manag 26:88–109. https://doi.org/10.1006/jeem.1994.1006

Schuman H, Presser S, Ludwig J (1981) Context effects on survey responses to questions about abortion. Public Opin Q 45:216–223. https://doi.org/10.1086/268652

Silva A, Nayga RM, Campbell BL, Park JL (2011) Revisiting cheap talk with new evidence from a field experiment. J Agric Resource Econ 36:280–291. https://doi.org/10.22004/ag.econ.117168

Stevens TH, Tabatabaei M, Lass D (2013) Oaths and hypothetical bias. J Environ Manage 127:135–141. https://doi.org/10.1016/j.jenvman.2013.04.038

Strange N, Jacobsen JB, Thorsen BJ, Tarp P (2007) Value for money: protecting endangered species on Danish Heathland. Environ Manage 40:761–774. https://doi.org/10.1007/s00267-006-0221-y

Strazzera E, Genius M, Scarpa R, Hutchinson G (2003) The effect of protest votes on the estimates of wtp for use values of recreational sites. Environ Resource Econ 25:461–476. https://doi.org/10.1023/A:1025098431440

Svenningsen LS (2019) Social preferences for distributive outcomes of climate policy. Climatic Change 157:319–336. https://doi.org/10.1007/s10584-019-02546-y

Talpur MA, Koetse MJ, Brouwer R (2018) Accounting for implicit and explicit payment vehicles in a discrete choice experiment. J Environ Econ Policy 7:363–385. https://doi.org/10.1080/21606544.2018.1450789

Tonsor GT, Shupp RS (2011) cheap talk scripts and online choice experiments: "looking beyond the mean." Am J Agric Econ 93:1015–1031. https://doi.org/10.1093/ajae/aar036

Tourangeau R, Rasinski KA (1988) Cognitive processes underlying context effects in attitude measurement. Psychol Bull 103:299–314. https://doi.org/10.1037/0033-2909.103.3.299

Tourangeau R, Rips LJ, Rasinski K (2000) The psychology of survey response. Cambridge University Press, New York

Uggeldahl KC (2018) Essays on decision making processes, information acquisition, and preferences in stated choice experiments: applications to economic valuation of consumer and environmental goods. https://ifro.ku.dk/english/research/past_phd_defences/2018/phd-23-august-2018/. Accessed 12 Aug 2019

Varela E, Mahieu P-A, Giergiczny M et al (2014) Testing the single opt-out reminder in choice experiments: an application to fuel break management in Spain. J Forest Econ 20:212–222. https://doi.org/10.1016/j.jfe.2014.05.001

Vedel SE, Jacobsen JB, Thorsen BJ (2015a) Contracts for afforestation and the role of monitoring for landowners' willingness to accept. Forest Policy Econ 51:29–37. https://doi.org/10.1016/j.forpol.2014.11.007

Vedel SE, Jacobsen JB, Thorsen BJ (2015b) Forest owners' willingness to accept contracts for ecosystem service provision is sensitive to additionality. Ecol Econ 113:15–24. https://doi.org/10.1016/j.ecolecon.2015.02.014

Von Haefen RH, Massey DM, Adamowicz WL (2005) Serial nonparticipation in repeated discrete choice models. Am J Agr Econ 87:1061–1076. https://doi.org/10.1111/j.1467-8276.2005.00794.x

Vossler CA, Doyon M, Rondeau D (2012) Truth in consequentiality: theory and field evidence on discrete choice experiments. Am Econ J Microecon 4:145–171. https://doi.org/10.1257/mic.4.4.145

Vossler CA, Evans MF (2009) Bridging the gap between the field and the lab: environmental goods, policy maker input, and consequentiality. J Environ Econ Manag 58:338–345. https://doi.org/10.1016/j.jeem.2009.04.007

Vossler CA, Zawojska E (2020) Behavioral drivers or economic incentives? Toward a better understanding of elicitation effects in stated preference studies. J Assoc Environ Resource Econ 7:279–303. https://doi.org/10.1086/706645

Whitehead JC, Blomquist GC (1999) Do reminders of substitutes and budget constraints influence contingent valuation estimates? Reply to another comment. Land Economics 75:483–484. https://doi.org/10.2307/3147193

Willis GB (2005) Cognitive interviewing: a tool for improving questionnaire design. Sage, Thousand Oaks, CA

Wright JD, Marsden PV (eds) (2010) Handbook of survey research, 2nd Revised. Emerald Group Publishing, West Yorkshire, England

Zawojska E, Bartczak A, Czajkowski M (2019) Disentangling the effects of policy and payment consequentiality and risk attitudes on stated preferences. J Environ Econ Manag 93:63–84. https://doi.org/10.1016/j.jeem.2018.11.007

Zawojska E, Czajkowski M (2017) Re-examining empirical evidence on stated preferences: importance of incentive compatibility. J Environ Econ Policy 6:374–403. https://doi.org/10.1080/21606544.2017.1322537

Zillmann D, Schmitz A, Skopek J, Blossfeld H-P (2014) Survey topic and unit nonresponse. Qual Quant 48:2069–2088. https://doi.org/10.1007/s11135-013-9880-y

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.