A fine balance:
specialized questioning techniques and their use in conservation

Jacopo Cerri[1], Elizabeth Oneita Davis[2], Diogo Veríssimo[2,3], and Jenny Anne Glikman[4]

[1]Dipartimento di Scienze della Vita e Biologia dei Sistemi, Università degli Studi di Torino, Via Accademia Albertina 13, 10123, Torino. **email**: jacopocerri@gmail.com

[2]San Diego Zoo Institute for Conservation Research, 15600 San Pasqual Valley Rd, Escondido, CA 92026

[3]Department of Zoology, University of Oxford, Oxford, UK

[4]Instituto de Estudios Sociales Avanzados (IESA-CSIC), Campo Santo de los Mártires 7, Córdoba, Spain

Warning

This is a ***preprint***, not a peer-reviewed study. If you do not know what a preprint is, we encourage you to read more about this type of documents (https://en.wikipedia.org/wiki/Preprint), before evaluating and citing the study.

Citing this preprint

On the 19th of March 2021, this preprint was published on Biological Conservation. Please cite it as: "Cerri, J., Davis, E. O., Veríssimo, D., Glikman, J. A. (2021). Specialized questioning techniques and their use in conservation: A review of available tools, with a focus on methodological advances. Biological Conservation, 257, 109089. https://doi.org/10.1016/j.biocon.2021.109089"

Abstract

Conservationists measuring noncompliance with rules about the exploitation of natural resources often need to ask sensitive questions. However, respondents can introduce bias through distorting their answers to direct questions, due to social norms and/or the risk of legal sanctions. Specialized Questioning Techniques (SQTs) are often a more suitable approach to counteracting respondent bias, as they encourage honest answering by protecting respondent's privacy by design. This study aims to provide a complete overview of the main SQTs, as well as about their most recent advances.

We performed a scoping review of existing SQTs, starting from those covered in Nuno and St. John (2015). We included techniques that have never been explained to conservationists before, and/or that were invented after 2015, or which had advanced and improved since Nuno and St. John (2015).

Our review identified 9 different broad types of SQTs, defined according to their practical implementation. We found 18 new versions of the randomized response technique, the unmatched count technique, the item sum technique, the triangular model and the crosswise model. We also discuss endorsement experiments and the ballot box method, which have not been covered

in a review for the benefit of conservationists. Finally, we found four new SQTs: the parallel model, the pair method, the list method and the person count technique. Each technique is explained with real, or hypothetical, examples and discussed in its advantages and limitations.

SQTs have undergone an impressive development over the last few years, and many different techniques are available to elicit sensitive behaviors with implications for conservation. This research offers a summary to conservationists and practitioners who want to understand SQTs and integrate them in their work, which should be a priority for those studying sensitive behaviors with implications for conservation.

1 Introduction

Conservation research often deals with human behaviors that are sensitive (Krumpal, 2013). These can include illegal activities subjected to formal sanctions, like poaching (St. John et al., 2015), taboos imposed by traditional beliefs (Jones et al., 2008), or socially undesirable activities that would meet with disapproval by our own peers.

As a result, conservationists who attempt to study the prevalence and drivers of such behaviors through questionnaire surveys may find that data gathered from respondents are biased (e.g. Davis et al., 2019; St. John, Mai, & Pei 2015). This bias often lies in the way answers are collected, which is usually through direct questioning, and has lower perceived privacy protection. When respondents are recruited in the field, they are (ideally) assured that their privacy is protected by researchers' commitment to ensure confidentiality of the findings. However, this commitment is often not enough to outweigh the potential costs connected with revealing deviant or illegal behavior and, even in confidential direct answers, respondents might feel their privacy is not adequately protected (Brittain et al., 2020). Moreover, in contexts characterized by human rights abuses, where local authorities may sift through researchers' data (Russo & Strazzari, 2020), direct answers could jeopardize both researchers and participants, especially for studies exploring deviant behaviors.

Specialized Questioning Techniques (SQTs) have been proposed as a way to overcome these issues, through anonymizing answers by design, and began to be adopted more widely in conservation almost fifteen years ago (e.g. Nuno, Bunnefeld, Naiman, & Milner-Gulland 2013; Solomon, Jacobson, Wald, & Gavin 2007). This increase in the use of SQTs was discussed in a systematic review in 2015 (Nuno & St. John, 2015), which provided researchers with a broad overview of existing methods and showed that Randomized Response Technique (RRT) and Unmatched Count Technique (UCT) were most commonly used; however, the authors noted that SQTs were still under-utilized in studies designed to accurately understand human behavior. Following this review, conservationists, albeit still adopting the classical designs of RRT or the UCT, were increasingly aware of their potential limitations (Davis et al., 2019; Hinsley, Keane, St. John, Ibbett, & Nuno 2019).

Many things have changed since 2015, in the field of SQTs. There has been a notable increase in the use of the "classical" SQTs of RRT and UCT (e.g. Cerri, Mori, Vivarelli, & Zaccaroni 2017; Cerri, Ciappelli, Lenuzza, Nocita, & Zaccaroni 2018; Chang, Cruyff, & Giam 2018; Davis et al., 2019; Davis and Glikman, 2020; Hinsley, Nuno, Ridout, St. John, & Roberts 2017; Ibbett et al., 2017; Ruppert et al., 2020), with corresponding advancements in understanding how useful these methods are in varying contexts and for varying behaviors. This is largely the result of continually increasing recognition of the importance of understanding and measuring conservation-related human behavior (Bennett et al., 2017), as well as deviant behavior in general (Gino & Ariely, 2016); for example, SQTs are now seen as a means to accurately measure prevalence to inform impact evaluations of behavior change campaigns (e.g. Davis et al., 2020), or to measure changes wrought by other initiatives such as community engagement and awareness-building (Ruppert et al., 2020). Moreover, ongoing global digitalization offered new fields of application for SQTs, and social scientists started designing methods suitable for Internet users. However, researchers became increasingly aware that SQTs are cognitively demanding, can suffer from large estimation errors, do not always remove response bias, and in many cases, do not allow researchers to link behaviors to their drivers (e.g. Chuang, Dupas, Huillery, & Seban 2019).

Despite these issues, it is undoubtedly important that researchers utilize methods that will ensure they gather the most accurate estimations of behaviors that can threaten natural resources. By doing so, researchers can more effectively advise on conservation management priorities and where best to allocate money and time, while also providing robust, evidence-based baselines that conservation interventions can be evaluated against.

To understand and illuminate potential opportunities for expansion into more effective SQTs in conservation, we reviewed the available social science and statistical literature about advances in SQTs since Nuno and St. John's (2015) review. Here, we investigate published critiques and methodological analyses of each technique, and discuss the main challenges that SQTs face. Our overall aim for this study is to: (*i*) explain existing SQTs, (*ii*) summarize the most recent advances in existing SQTs and (*iii*) introduce novel SQTs that appeared after 2015, while providing practical examples for practitioners and researchers working in conservation who have used SQTs before, or plan to use them for the first time, and illuminating potential limitations and strengths of these methods.

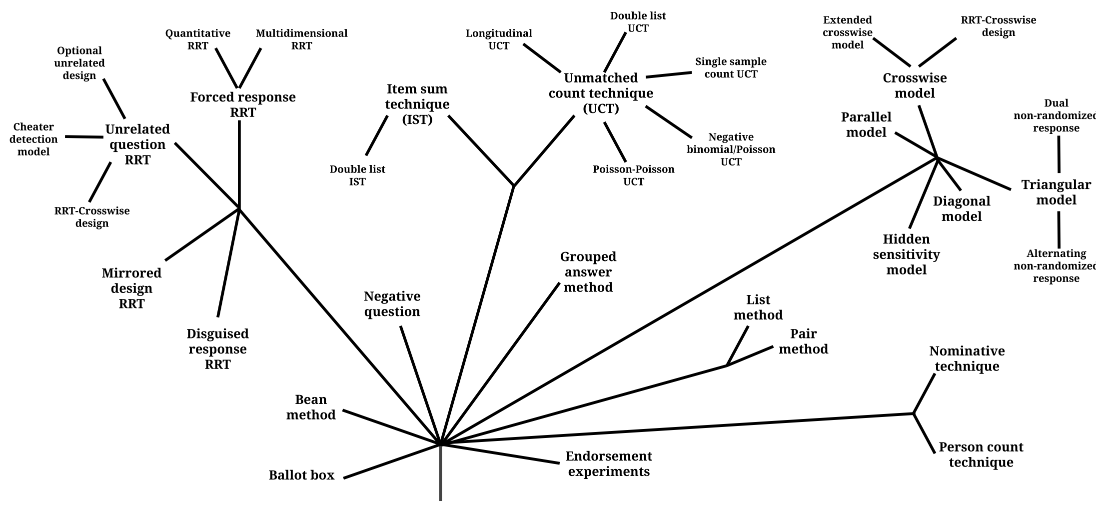

Figure 1 | Groups of SQTs, single techniques and their advances.

2 Methods

We conducted a scoping review of SQTs. The adoption of a scoping review was motivated by the fact that, when replicating the query used in 2015 (Nuno & St. John 2015) on ISI and Scopus, we found more than 9,000 studies, often not directly related to deviant or illegal behavior. Most of these studies were adopting questionnaires for diverse reasons, and only marginally mentioned social desirability or response bias in the discussion. The effectiveness of the original query was probably affected by growth in indexed scientific publications in the English language, which almost doubled since 2014 (Bornmann & Mutz, 2015).

Our scoping review was nevertheless based on a structured search protocol. First, starting from Nuno and St. John (2015), we identified 9 SQTs, each one characterized by one, or more, seminal studies: the randomized response technique (RRT: Greenberg, Abul-Ela, Simmons, & Horovitz 1969; Kuk, 1990; Warner, 1965), the unmatched count technique (UCT, also known as "list experiment", Droitcour et al., 2004), the item sum technique (Trappmann, Krumpal, Kirchner, & Jann 2014) the nominative technique (Miller, 1985), the grouped-answer method (also known as the "three-card method", Droitcur & Larson, 2002), the crosswise model (Yu, Tian, & Tang 2008), the triangular model (Yu, Tian, & Tang 2008), the diagonal model (Groenitz, 2014), the hidden sensitivity model (Tian, Yu, Tang, & Geng 2007), the negative question method (Esponda & Guerrero, 2009), and the bean method (Lau, Yeung, Mui, Tsui, & Gui 2011). For each one of these 9 SQTs we read those studies that had proposed them for the first time, and then we searched on Google Scholar for those studies published after 2015, which cited them. To increase the effectiveness of our strategy we also searched for those studies that cited two major reviews about SQTs in conservation (Hinsley, Keane, St. John, Ibbett, & Nuno 2019; Nuno & St. John, 2015).

In addition, we also considered SQTs which appeared after 2015 or were not covered in Nuno and St. John (2015), and which we had encountered during our on-line search of related articles: the person count technique, the parallel model, the list method, the pair method, the ballot box method and endorsement experiments.

We grouped SQTs based on differences in their practical implementation (e.g. methods using a randomizing device), rather than according to differences in the statistical estimation of parameters (e.g. the crosswise model can be regarded as a variant of the RRT). We deem this classification to be more helpful for conservation scientists, to better grasp the underlying idea behind the various SQTs and to imagine their practical application in various settings.

3 Results

In terms of their privacy protection mechanism, we found 9 main groups of techniques (Table 1). The nominative technique, the grouped-answer method, the diagonal model, the hidden sensitivity model, negative questions and the bean method did not receive any significant advance in their practical implementation, since 2015, although some of these methods have now begun to be adopted by conservation researchers (e.g. the bean method, Cerri, Ciappelli, Lenuzza, Nocita, & Zaccaroni 2018; Jones, Papworth, Keane, Vickery, & St. John 2020). Since 2015, we found 4 invented new SQTs and 18 new variants of existing techniques (Table 1). Below we briefly explain these methods and show an approximate "phylogenetic" tree of the connections between the various methods (Fig. 1)

3.1 A short overview of existing SQTs

In the following lines, we aim to provide an introduction to the three main types of SQTs, which were advanced since 2015. Notably, we will discuss the randomized response technique, the unmatched count technique, the item-sum technique, and two non-randomized techniques: the crosswise and the triangular model.

3.1.1 The randomized response technique

The Randomized Response Technique (RRT) introduces some noise, with a known probability distribution, to mask individual answers (Nuno & St. John, 2015). Various statistical approaches are then adopted to handle this noise, estimate behavioral prevalence and also obtain likelihood functions to model individual covariates (Cruyff, Böckenholt, Van Der Heijden, & Frank 2016).

Noise can come from a randomizing device or from one or more non-sensitive questions. The higher the amount of noise, the higher the level of privacy protection, but at the cost of the statistical efficiency of the estimator; moreover, simple probabilities are easily calculated and can incur suspicion in respondents if they believe the researcher can record whether they answered truthfully or not.

However, it is possible to play with people's misperception of probabilities to increase the efficiency of the randomization (e.g. by using the "Benford illusion", where people misperceive the real distribution of the first digits of a numeric series, Diekmann, 2012). In terms of their practical implementation, there are four major approaches to the RRT. Importantly, all of them assume that respondents are left alone when answering the question, and that their use of the randomizing device is hidden from the interviewer (e.g. with a cup over a die).

In the forced response design (Boruch, 1971), the most common RRT variant in conservation studies (e.g. Cerri, Mori, Vivarelli, & Zaccaroni 2017; Davis et al., 2019; Ruppert et al., 2020; Santangeli, Arkumarev, Rust, & Girardello 2016; St. John, Mai, & Pei 2015), respondents use the randomizing device to ensure that it is impossible for researchers to know which options were selected in a multi-option question, or which numerical answers were provided. For example, respondents may be given a traditional six-sided die. If respondents roll 1, they are asked to give a forced response of "no" to the question. If they roll 6, they must give a forced response of "yes". If they roll 2, 3, 4, 5 then they can answer truthfully (Fig. 2). The biggest advantages of the forced-response RRT are: that its only assumptions are respondent privacy being maintained by the researcher stepping away and/or covering the randomizing device, along with respondents' compliance with instructions; and statistical efficiency, which is higher than the other RRTs. However, it requires the use of a randomizing device, which has three main limitations: (*i*) respondents must understand how to use it, (*ii*) respondents should be familiar with that particular device, to understand how it will protects their privacy and (*iii*) the context where the questionnaire is administered should be suitable for using the randomizing device (e.g. a die should not be used if gambling is frowned upon in the society the researcher is working in).Finally, although statistical error may not be as large as techniques like unmatched count technique (UCT) (Section 3.1.2), the errors can still be large enough to make accurate estimations of prevalence challenging (e.g. Ruppert et al., 2020)

In the unrelated question design (Greenberg, Abul-Ela, Simmons, & Horovitz 1969), the randomizing devices allocate respondents to one of two questions, with identical response options (e.g. Yes/No). Privacy is therefore protected because it is impossible to understand to which question respondents were allocated (e.g. Lee, Peng, Tapsoba, & Hsieh 2017). One of the two questions is the sensitive question of interest, while the other is a non-sensitive question, whose prevalence in the target population is known (Fig. 3).

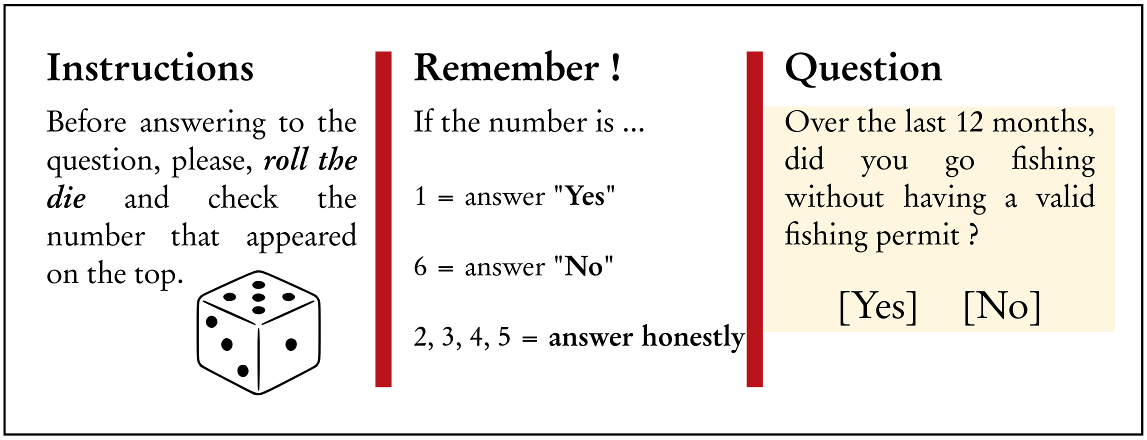

Figure 2 | An example instruction card for the forced-response RRT, containing a case study about illegal fishing (Cerri, Ciappelli, Lenuzza, Zaccaroni, & Nocita 2018). Respondents are provided with a 6-faces die, instructions and the question. The die must be rolled before answering and respondents must not reveal the outcome of die roll. Respondents could record their answers on a self-administered questionnaire (like in this example) or answer to an enumerator.

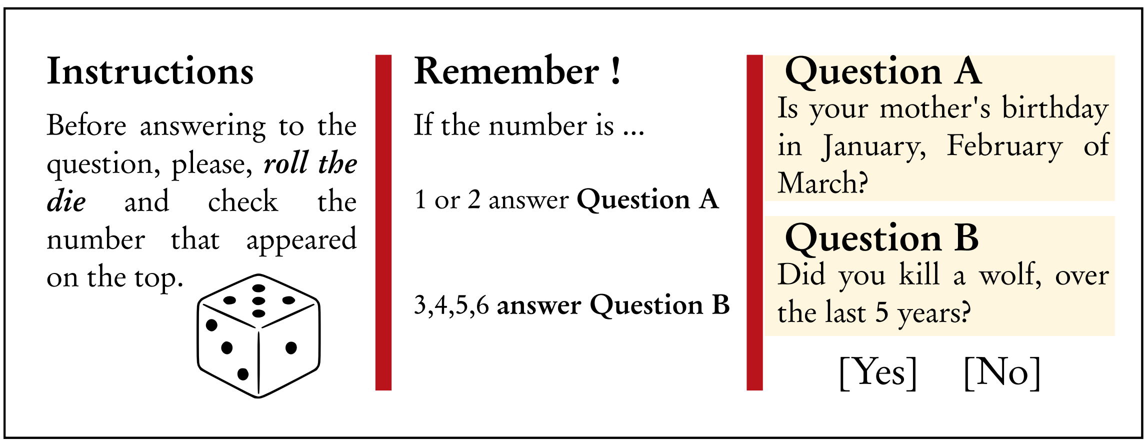

Figure 3 | An example instruction card for the unrelated question RRT, containing a hypothetical case study about wolf killing. Respondents are provided with a 6-faces die, instructions and the question. The die must be rolled before answering and respondents must not reveal the outcome of die roll. Respondents could record their answers on a self-administered questionnaire (like in this example), or answer to an enumerator.

In the mirrored question design (derived from the indirect questioning mode, Warner, 1965), respondents are asked one of two mirrored sensitive questions/statements, according to probability (Blair, Imai, & Zhou 2015). Like in the unrelated question design, here the randomizer allocates respondents to one of two conditions, thus making it impossible to understand which question respondents were allocated to (Fig. 4).

In the disguised response design (Kuk, 1990), respondents never need to state “true/yes” or “false/no” to the interviewer. Instead, they are given two decks, which correspond to the real answer of respondents (one deck for “true-yes” and one deck for “true-no”) with different proportions of colored cards, corresponding to “stated” yes/no. For example, Chang, Cruyff, and Giam (2018), in a study investigating bird hunting in China provided respondents with two decks with 15 cards each, one for the real “yes” and one for the real “no” answer. The “yes deck” had 80% red and 20% black cards, while the “no deck” contained 20% red and 80% black cards.

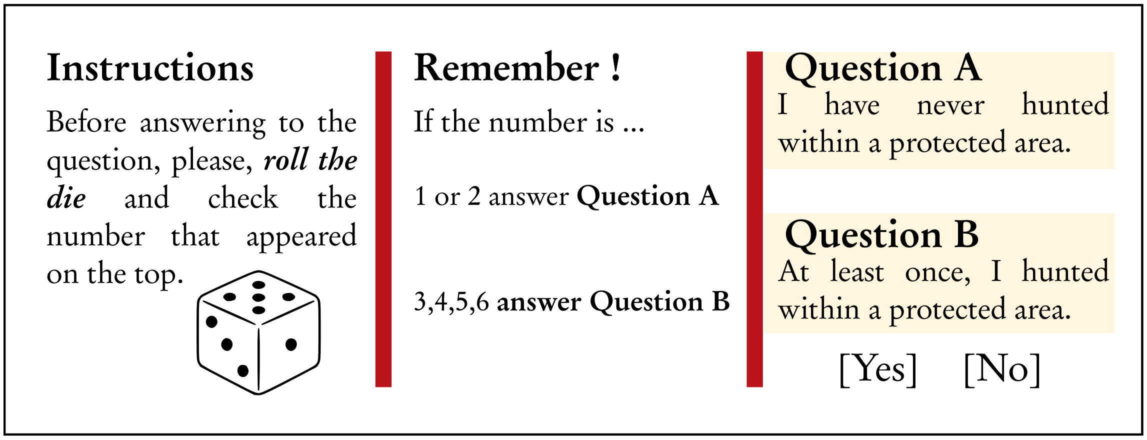

Figure 4 | An example instruction card for the mirrored question RRT, containing a hypothetical case study about illegal hunting. Respondents are provided with a 6-faces die, instructions and the question. The die must be rolled before answering and respondents must not reveal the outcome of die roll. Respondents could record their answers on a self-administered questionnaire (like in this example), or answer to an enumerator.

Red cards corresponded to a "yes" response and black card with a "no" response. Respondents were asked to: (*i*) draw a card from the deck corresponding to their true answer (e.g. from the "yes" deck if they really engaged in the behavior), (*ii*) not reveal which deck the card was from, (*iii*) report the card's color, (*iv*) put the card back into the deck and (*v*) shuffle the deck (Fig. 5). In the disguised response the privacy is protected by the fact that researchers do not know from which deck respondents drew the card. Chang, Cruyff, and Giam (2018) found that respondents were more likely to believe that disguised response gave them anonymity, compared to the forced-response. This indicates that disguised response may be effective at increasing respondents' comfort and therefore their compliance with instructions, compared to other RRTs.

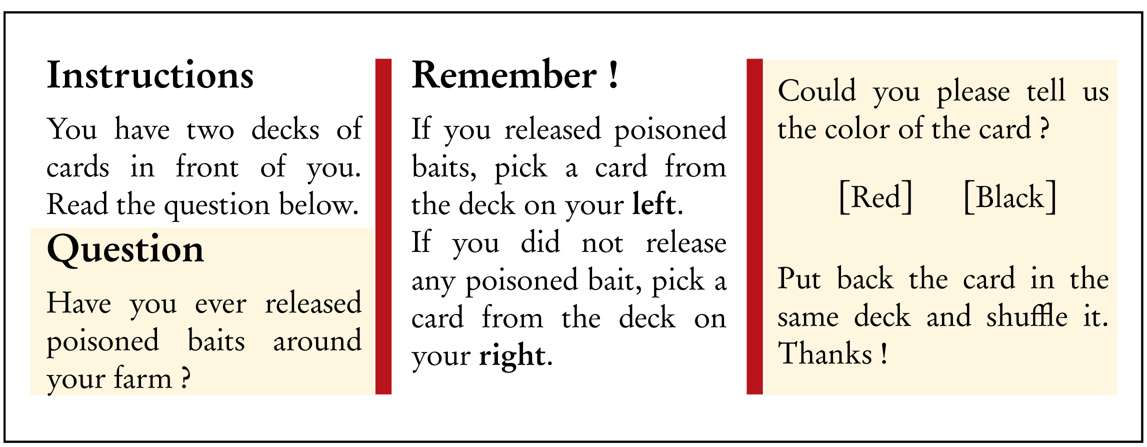

Figure 5 | An example instruction card for the disguised response RRT, containing a hypothetical case study about illegal hunting. Respondents are provided with 2 decks of cards on a table, instructions and the question. A card must be drawn before answering and respondents must not reveal from which deck it was extracted. Respondents could record their answers on a self-administered questionnaire (like in this example), or answer to an enumerator.

3.1.2 The unmatched count technique (list experiments) and the item sum technique

The Unmatched Count Technique (hereafter UCT, also known as “list experiment”) protects respondents’ privacy by asking them simply to indicate how many behaviors apply to them, from a list of non-sensitive and sensitive behaviors. In its classical version, the item count technique (Raghavarao & Federer, 1979) respondents are separated into two groups, one designated as the control group and one as the treatment. Each group is shown a set of behaviors and asked how many of these they have done. The control group is only shown innocuous behaviors, while the treatment group is shown the same set of innocuous behaviors plus the sensitive behavior of interest (Fig. 6). The difference between the control and the treatment groups represents the mean prevalence of the target sensitive behavior in the population. The item count technique is argued to reduce sensitivity by having respondents only give a numeric response; thus, the researcher never knows if one of the behaviors performed by the respondent was the sensitive one of interest. Although this method requires significant checking and piloting of the behaviors utilized (Hinsley, Keane, St. John, Ibbett, & Nuno 2019), it is also simple to check for internal consistency (Blair & Imai, 2012). For example, researchers should ensure that there are no “floor effects”, where individuals would have to state “0”. Therefore, if the researcher finds multiple “0s” within their sample, they know that respondents were actively lying (e.g. Hinsley, Nuno, Ridout, St. John, & Roberts 2017). A major advantage of the UCT lies in the trust it inspires in respondents: Hoffmann, De Puiseau, Schmidt, and Musch (2017) found that the UCT was perceived as more capable of protecting respondents’ privacy then the crosswise method, the stochastic lie detector and the cheating detection model. Therefore, this method was considered to be particularly suitable for eliciting extremely sensitive behaviors, although this could be context dependent; a study conducted in Cambodia found that RRT was believed to be more trustworthy compared to UCT (Davis et al., 2019).

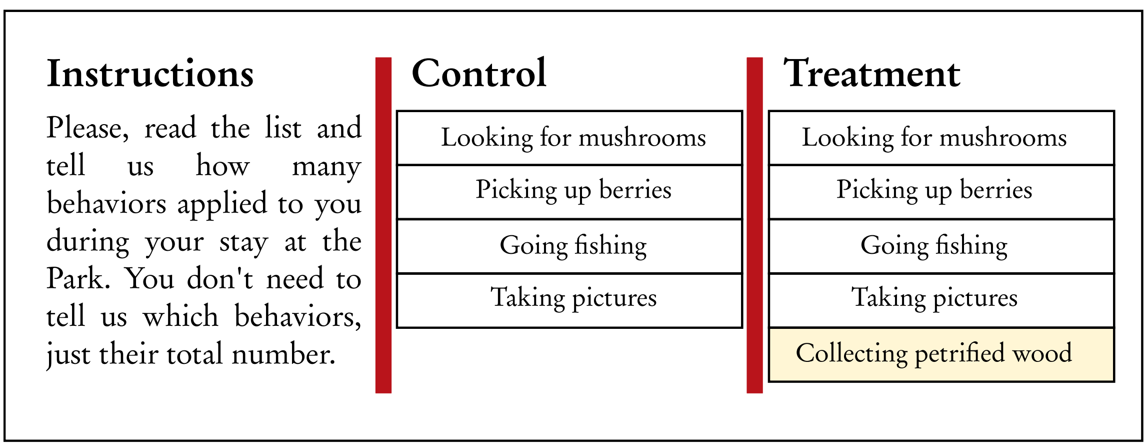

Figure 6 | An example instruction card for the classic 2-groups item count technique, containing a hypothetical case study about illegal petrified wood removal from a protected area (Cialdini et al., 2006). Respondents are randomly allocated to the Control or Treatment condition and should write on the questionnaire, or tell the interviewer the number of actions they did, during their stay at the Park. The sensitive behavior is highlighted.

In double list UCT, the respondents are further separated, with two sets of non-sensitive behaviors, A and B, and the appended sensitive item switched off between the lists (Glynn 2013, Fig. 7). This method can in principle narrow the large statistical errors that characterize UCT (e.g. Davis et al., 2019, 2020; Hinsley, Keane, St. John, Ibbett, & Nuno 2019). However, to our knowledge it has not been applied in a conservation setting

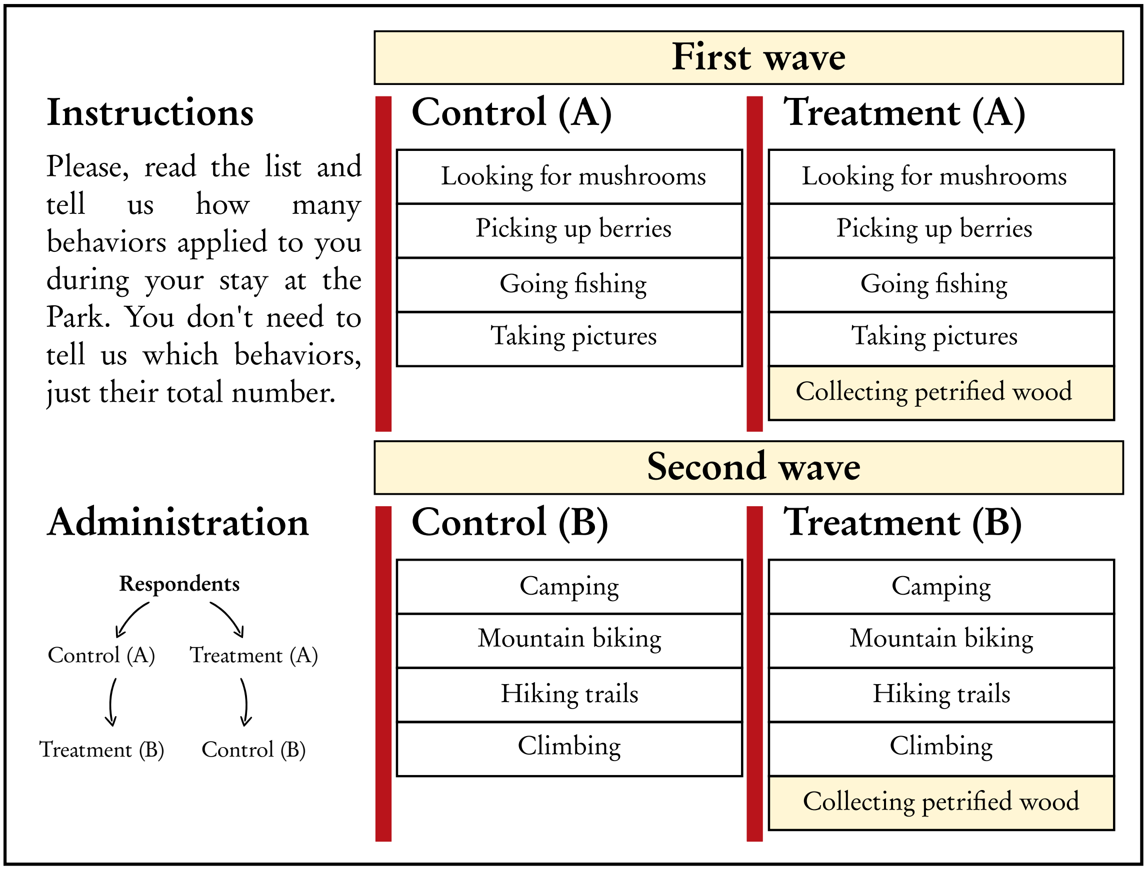

Figure 7 | An example instruction card for the double list unmatched count technique, containing a hypothetical case study about illegal petrified wood removal from a protected area (Cialdini et al., 2006). Respondents are randomly allocated to the Control or Treatment condition in the first wave, and should write on the questionnaire, or tell the interviewer the number of actions they did, during their stay at the Park. Then on the second wave, they are allocated to the opposite condition. Control lists in the first and second wave contain two different sets of behaviors. The sensitive behavior is highlighted.

In addition, where the perceived sensitivity of certain behaviors is unknown, the design of the double list UCT can provide some measure of this sensitivity. For example, in their study of women's sexual behaviors in Côte d'Ivoire, Chuang, Dupas, Huillery, and Seban (2019) crafted multiple lists of 48 statements total, divided into two sets. In their "treatment" set, the lists included innocuous as well as semi-sensitive statements, e.g. "Many women have an abortion even though it is illegal". They could therefore compare statements between their control set to see whether women generally answered at a lower prevalence for any perceived sensitive statements. However, this consistency and sensitivity check comes with risk to the respondents. Ceiling effects occur when a respondent states that they have done and/or agree with the statements presented, thus negating their ability to give an anonymous response. Chuang, Dupas, Huillery, and Seban (2019) found that ceiling effects occurred as affirmative responses grouped, e.g. women who stated yes to a sexual behavior were also more likely to state yes to questions such as "Many women have an abortion even though it is illegal".

In single sample count UCT, the sample is not split into two groups. Instead, the entire sample is asked about the sensitive behavior, as well as being asked four non-sensitive questions where each of the questions must have a 50% yes rate (Petróczi et al., 2011, Fig. 8). Thus, the estimated prevalence of the behavior concerned will be any deviation from 50%. Although this method was conceived to overcome the large standard errors of UCT, the method still suffered from this issue. Similarly to double list UCT, there have been no attempts in the conservation literature to apply this method.

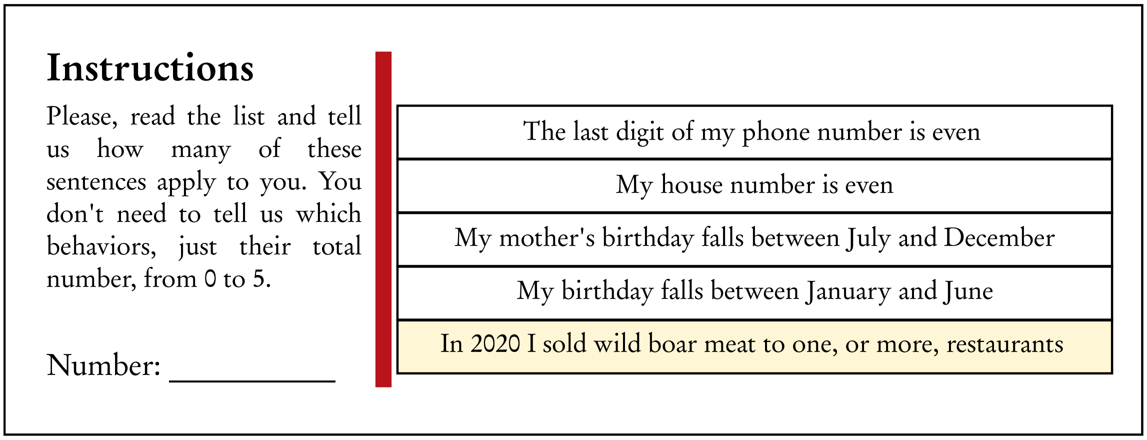

Figure 8 | An example instruction card for the single sample count unmatched count technique. Respondents here might be hunters, in a hypothetical study where researchers want to estimate how many of them illegally sold wild boar meat to one, or more, restaurants in 2020. Answers could be recorded directly on the questionnaire (like here), or can be dictated to an interviewer. The sensitive behavior is highlighted.

Finally, the item sum technique (IST) is an extension of the UCT where respondents are asked to give summed quantitative answers, e.g. average number of wolves shot by hunters (Fig. 9).

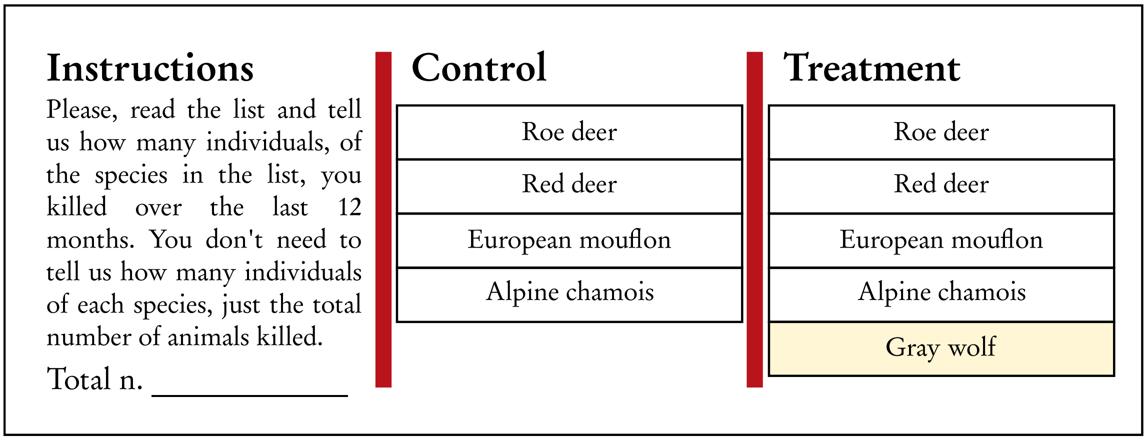

Figure 9 | An example instruction card for the item sum technique. Respondents here might be deer stalkers in the Alps, in a hypothetical study where researchers want to estimate their illegal killing of wolves. Answers could be recorded directly on the questionnaire (like here), or can be dictated to an interviewer. The sensitive behavior is highlighted.

3.1.3 Non-randomized techniques

Non-randomized techniques ask respondents to provide a conjoint answer to two or more questions, including the sensitive question of interest. They do not require a randomizing device, and therefore their field administration is made easier than RRT. Moreover, some of them, like the crosswise model, ensure that it is difficult for respondents to adopt self-protective answers.

In the crosswise model, respondents are asked only to tell whether they provided the same answer, or two different answers, to a couple of statements. One statement is the non-sensitive question, whose prevalence in the target population must be known in advance and the second statement is about the sensitive behavior of interest (Fig. 10).

Instructions

Please, read Statement A and Statement B. Then, answer the question.

Statement A

My mother was born between August and December.

Statement B

I consumed pangolin meat, at least once.

Question

Are both Statement A and Statement B, true or false (you would have answered True/True and False/False to them) ?

[Yes] [No]

Figure 10 | An example instruction card for the crosswise method, from a hypothetical study exploring the consumption of Pangolin meat in South-East Asia. Answers could be recorded directly on the questionnaire (like here), or can be dictated to an interviewer. Of course, researchers must know the distribution of Question A, to estimate question B.

Compared to direct questioning and the UCT, Hoffmann, De Puiseau, Schmidt, and Musch (2017) found that the crosswise model was easy to understand for respondents', and that the method was deemed to be good at protecting individual privacy, being outperformed only by the UCT. Furthermore, Meisters, Hoffmann, and Musch (2020) found that providing respondents with detailed instructions and additional comprehension checks might improve its accuracy. The performances of the crosswise model are still debated: some authors found that the technique is sensitive to false positives (i.e., where respondents state that they have performed the sensitive behavior when they have not (Höglinger & Diekmann, 2017); however, validation studies based on experimentally induced cheating behavior (Hoffmann, Diedenhofen, Verschuere, & Musch 2015) and known prevalence of sensitive behaviors (Korndörfer, Krumpal, & Schmukle 2014) supported the crosswise model's robustness as a specialized questioning technique.

The triangular model (Yu, Tian, & Tang 2008) is similar to the crosswise model in the sense that respondents are provided with a couple of questions: a safe one, with known prevalence and a sensitive one, about the sensitive trait which should be estimated. However, rather than indicating whether they provided the same answer or not, they are asked whether they answered "no" to both questions, or whether they affirmed one (Fig. 11). It is easy to see that the triangular model provides respondents with a clear self-protective strategy, as they could simply state that they answered "no" to both questions. Although the triangular method protects respondents' privacy more efficiently than the crosswise model, this can cause estimations to be incorrect because respondents are so trusting of the method that they trust that they can answer deceitfully and not be detected (this effect has also been proposed to occur in at least one case, with RRT (Davis et al., 2019). Overall, the crosswise model might be preferred, as it might be harder for respondents to violate instructions, as the crosswise model does not provide any clear self-protective answering behavior.

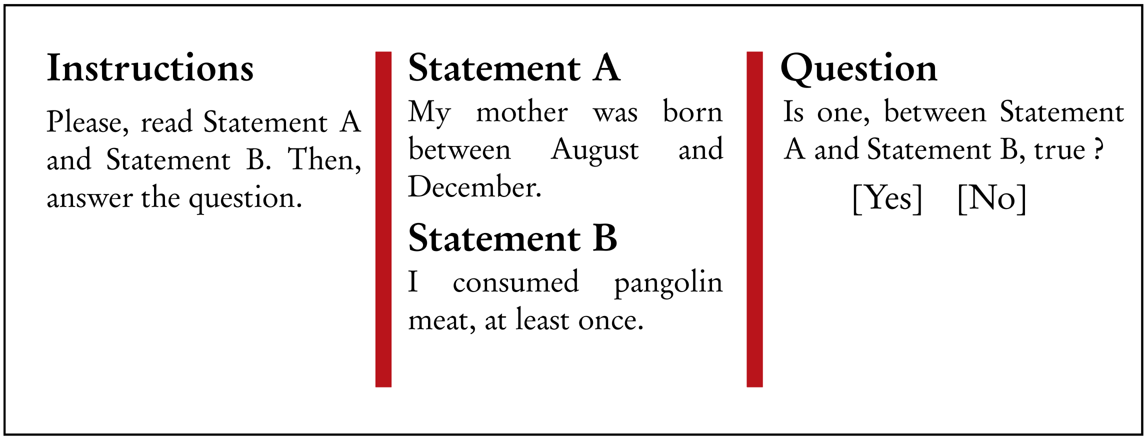

Figure 11 | An example instruction card for the triangular method, from a hypothetical study exploring the consumption of Pangolin meat in South-East Asia. Answers could be recorded directly on the questionnaire (like here), or can be dictated to an interviewer. Of course, researchers must know the distribution of Statement A, to estimate the frequency of Statement B.

3.2 Advances in existing techniques

3.2.1 Advances in RRT

The quantitative RRT, for Poisson-distributed count data (Cao, Breidt, Solomon, Conteh, & Gavin 2018) follows the same method as the forced response (Boruch, 1971), but the response is a count variable. A sealed vessel with colored balls can be used as a randomizing device. Balls have a number from a known Poisson distribution. Respondents draw a ball and, based on its color, report their count answer or the number on the ball (Fig. 12). Compared to additive or multiplicative RRT (Eichhorn & Hayre, 1983; Pollock & Bek, 1976), where respondents must engage in mathematics when giving their numerical response, the advantage of this approach lies in the fact that respondents do not have to sum or multiply quantities, thus reducing their cognitive load.

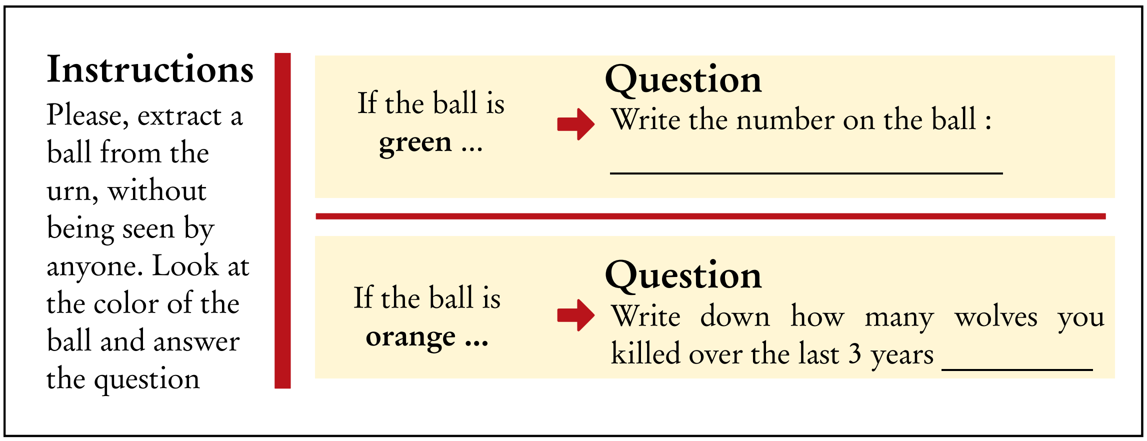

Figure 12 | An example instruction card for the quantitative RRT, containing a hypothetical case study about wolf persecution from herders. Before answering, respondents draw a ball from a sealed vessel. Balls are numbered with numbers from a known Poisson distribution. If the ball is green they are asked to write or state the number on the ball, while if the ball is orange they can write or state down how many wolves they killed. Respondents can also simply report a number to an interviewer. Privacy is protected if the color of the ball is seen by respondents only.

The multidimensional RRT (Cruyff, Böckenholt, & van der Heijden 2016) is based on two nested forced-response questions. The first question estimates whether respondents engaged in the target behavior, while the second question quantifies its frequency (Fig. 13).

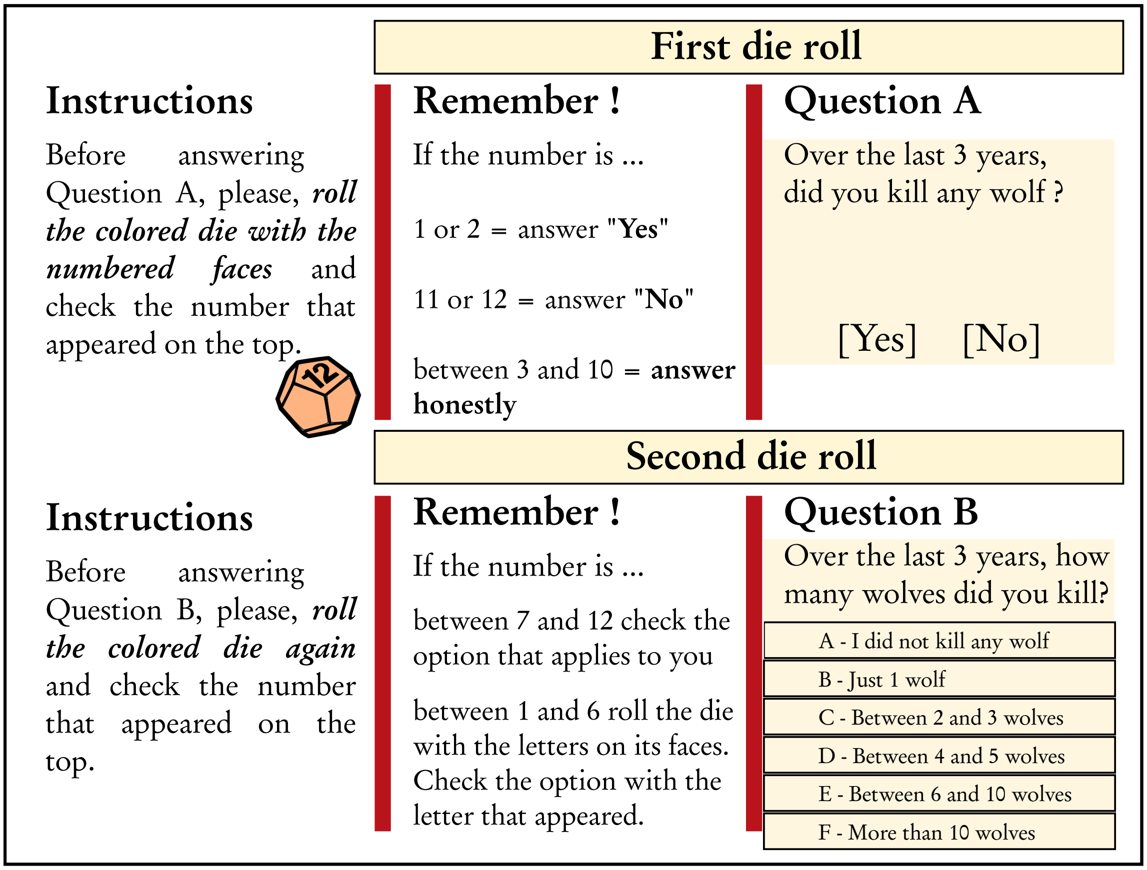

Figure 13 | An example instruction card for the multidimensional RRT, containing a hypothetical case study about wolf persecution from herders. Respondents are provided with a 12-faced colored die, with numbers on it, and with a 6-faces die with letters from A to F on its faces. They are asked to roll the die colored die first, answer to Question A, and then move to question B. The technique is mostly suitable for self-administered questionnaires.

The multidimensional design is more efficient than unidimensional designs, as it can detect and account for noncompliance with instructions, and it can be adapted for the estimation of multinomial outcomes (Cerri et al., 2018). However, the adoption of two consecutive forced-responses raises some doubts about its cognitive load and its practical application in many conservation settings, especially those characterized by low levels of literacy. It is worth noting that approaches to detect and account for non-compliance in unidimensional forced-response RRT have also been proposed (Blair, Imai, & Zhou 2015; Chuang et al., 2019).

With respect to the unrelated question design (Greenberg, Abul-Ela, Simmons, & Horovitz 1969), as a certain question might be sensitive for one respondent, but not for another, the optional unrelated question design RRT was invented. Respondents are given the option of using a randomization device to answer a question about a sensitive behavior, or to simply answer the question directly if they believe that it is not sensitive (Fig. 14). If respondents are allocated to question $1_{\text{sensitive}}$ and believe the topic not to be sensitive, they can just answer question $2_{\text{sensitive}}$ directly. Respondents are given this option through an initial randomization device; thus, and the researcher does not ideally know whether the respondent has chosen to use the second randomization device or not. This technique ensures the accurate estimation of target behaviors, even on an ordered scale, while obtaining an estimation of their perceived sensitivity (Arnab & Rueda, 2016). While this method was evaluated from a mathematical standpoint, there has to date been only one application in a real scenario (sexual abuse: Chhabra, Dass, & Gupta 2016). Just like the multidimensional RRT, this technique might be too cognitively demanding for respondents, as it combines two consecutive, interdependent questions. Moreover, it is unclear whether the method might really make researchers unaware about the usage of the randomizing device, which is the basic protection mechanism of the whole method.

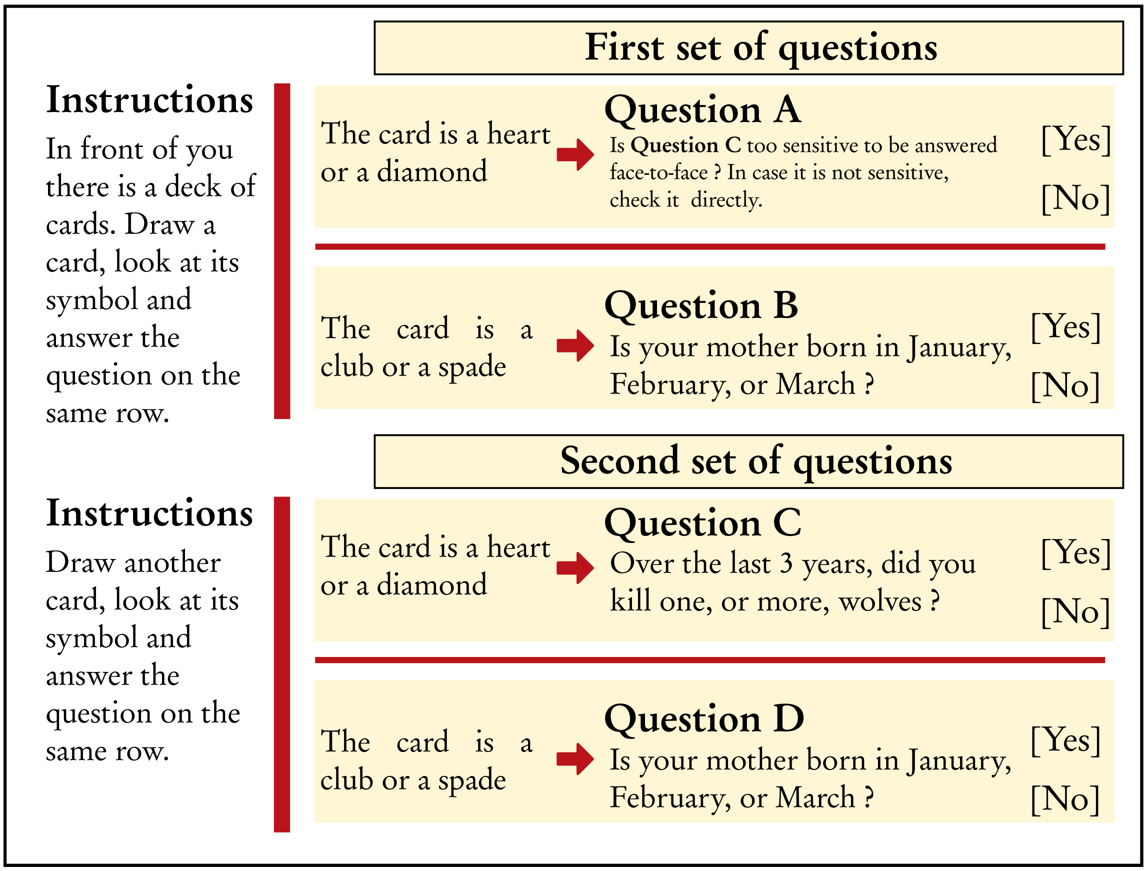

Figure 14 | An example instruction card for the optional unrelated question RRT, containing a hypothetical case study about wolf persecution from herders. Respondents are provided with a deck of cards, containing 60% of red cards (hearts or diamonds) and 40% of black cards (clubs or spades). This method is suitable for self-administered questionnaires only.

The cheater detection model (Clark & Desharnais, 1998) is also based on the unrelated question RRT. In the cheater detection model, there are two sub-samples with the neutral question as the "gatekeeper" (Fig. 15). This allows for the estimation of respondents who answered "no" to both questions. Although this can detect this "superficial" cheating, it's not possible to detect those who answer yes for the neutral question and then "no", even if they have in fact performed the sensitive behavior. Nonetheless, it provides a good baseline for the level of cheating occurring.

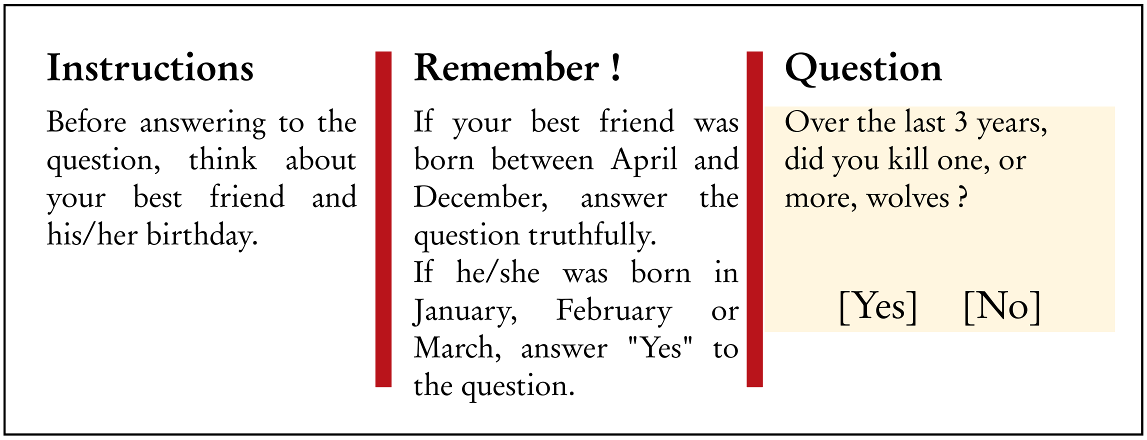

Figure 15 | An example instruction card for the cheater detection model, containing a hypothetical case study about wolf persecution from herders. Respondents could complete a self-administered form or answer to an interviewer.

The RRT/crosswise design, adopts the unrelated question with known prevalence to signal noncompliance. For example, if the unrelated question is “Is your age an even number?”, it is possible to check if the answers to this question align with the age stated within the survey. This can improve the estimation of sensitive attributes through a weighted conditional likelihood estimator (Tu & Hsieh, 2017). This technique has not been applied in a conservation setting.

Item response theory (IRT) has been proposed to estimate the prevalence of a sensitive behavior based on multiple RRT questions, while indicating the degree of noncompliance. Chang, Cruyff, and Giam (2018) applied IRT assessed to understanding the prevalence of illegal bird hunting by using three distinct RRT questions about three types of bird, e.g. “*Have you ever hunted partridge?*”, “*Have you ever hunted bulbuls?*”, etc. By summing the scores from these single RRTs, they were able to obtain an overall estimate of illegal bird hunting in their sample. In addition, as noncompliance might be signaled from common method bias (i.e. when they answered “no” to all three questions), the use of multiple RRT questions can indicate when a respondent is likely to have answered deceitfully. Although this type of control against bias requires additional analytical work for the researcher, this approach offers considerable opportunities for the analysis of RRT data while ensuring prevalence estimation accuracy.

Chuang, Dupas, Huillery, and Seban (2019) also developed internal consistency test to detect cheating in questionnaires with multiple RRT. For example, in a questionnaire asking respondents to flip a coin before answering each of multiple RRT, the expected frequency of “yes” and “no” answers should be around 0.5 and departures from this frequency signal noncompliance. Performing consistency tests is important for impact evaluation purposes, which SQTs are increasingly being used for; for example, non-compliance may actually be less prior to a campaign intended to reduce the sensitive behavior, and more after the campaign, due to norm pressure, thus obscuring accurate estimations of change (Camilotti, 2016). Although tests for internal consistency cannot directly measure such attributes, they can provide some indication of the trust researchers can have in their RRT estimations.

3.2.2 Advances in UCT and the item sum technique

In the longitudinal unmatched count technique (Gaia & Al Baghal, 2019) the same sample of respondents is given the control list and the sensitive list at different points in time. This method reduces the statistical error that UCT suffers from; however, the method requires a significant time investment, and as with UCT, if the non-sensitive items are not carefully chosen the method can still generate incorrect estimates. To date, the method has never been applied in for conservation issues.

In the negative binomial and Poisson UCT (Tian, Tang, Wu, & Liu 2017), respondents are divided into a control and a treatment group. Rather than asking respondents how many separate behaviors they have engaged in, respondents in the control group are asked to answer a non-sensitive count question, while those in the treatment condition to sum this answer with 1, in case they have engaged in a target sensitive behavior (Fig. 16). The respondent then states the sum, which anonymizes individual answers. While this method, altogether with Poisson and Negative Binomial models, can in principle overcome floor and ceiling effects, it introduces large variability in the data, which might jeopardize the accurate estimation of behavioral prevalence.

Instructions	Control	Treatment
Read the following questions and aswer.	How often have you visited Hanoi, over the last 12 months ? Report the number of times. N.visits: ____________	How often have you visited Hanoi, over the last 12 months? Add 1 to the number of visits to Hanoi, if you also used tiger bone glue, over the last 12 months N. visits: ___________

Figure 16 | An example instruction card for the negative binomial/Poisson count UCT, containing a hypothetical case study about tiger bone glue usage in Vietnam. Respondents are allocated at random to control or treatment groups.

The Poisson-Poisson UCT is an extension of the Poisson UCT, which aims to understand what covariates could affect the responses to the sensitive question of interest (Liu, Tian, Wu, & Tang 2019). Respondents are divided into two groups, with Group 1 asked to choose a random number from a Poisson distribution and report it. Group 2 is also asked to choose a random number from the Poisson distribution, but to sum this number with 1, if they have engaged in the sensitive behavior of interest (Fig. 17). Through this method, one can also collect covariate data such as age and gender, and explore the effect of these variables on response using a standard Poisson linear regression. However, the method is still characterized by highly variable estimates.

Instructions	Control	Treatment
Choose one of the following numbers. 6 8 4 4 7 4 2 3 6 5 1 5 2 2 4 2 5 10 3 5 3 3 3 3 4 4 5 3 5 4 8 5 2 3 6 7 13 6 5 1 5 7 4 5 7 8 3 4 2 7 1 2 4 7 5 4 6 5 2 4 5 4 4 4 7 4 5 8 6 6 6 3 8 4 6 5 8 10 11 1 3 7 8 7 8 1 4 5 7 10 6 6 4 48 2 5 7 3 8 5 4 6 2 10 9 3 6 7 6 10 7 3 4 8 6 4 5 6 6 3 4 5 6 8 9 1 4 4 5 8 7	Report the number you chose: ____________	If you used tiger bone glue, over the last 12 months, add 1 to the numer you chose. If you do not use tiger bone glue, report the number. Number: ____________

Figure 17 | An example instruction card for the Poisson-Poisson count UCT, containing a hypothetical case study about tiger bone glue usage in Vietnam. Respondents are allocated at random to control or treatment groups. Numbers on the list are extracted from a Poisson distribution with mean and variance equal to 5.

In the double list item sum technique (Krumpal, Jann, Korndörfer, & Schmukle 2018), the logic exactly follows the double list UCT. However, the double list item sum technique can easily be extended to include additional sensitive questions (e.g. Perri, Rodríguez, & García 2018). The double list item sum technique greatly reduces the variability intrinsic in standard item sum technique, but it might be cognitively demanding, and requires further testing.

3.2.3 Advances in non-randomized techniques

In the extended crosswise model (Heck, Hoffmann, & Moshagen, 2018), respondents are randomly assigned to two groups, as opposed to the crosswise model, where the sample is not split into multiple groups. The groups receive the same sensitive question, but two different, non-sensitive questions, with a known, and complementary prevalence in the target population (Fig. 18). If respondents in group A receive the question "*Is your mother's birthday between June and December?*", respondents in group B should receive the question "*Is your mother's birthday between January and May?*". As with the crosswise model, the respondent then states whether their answer is the same to both the non-sensitive and sensitive question, or whether the responses are different. The extended crosswise model, while guaranteeing the same statistical efficiency of the crosswise model, can also enable detection of respondent noncompliance with instructions.

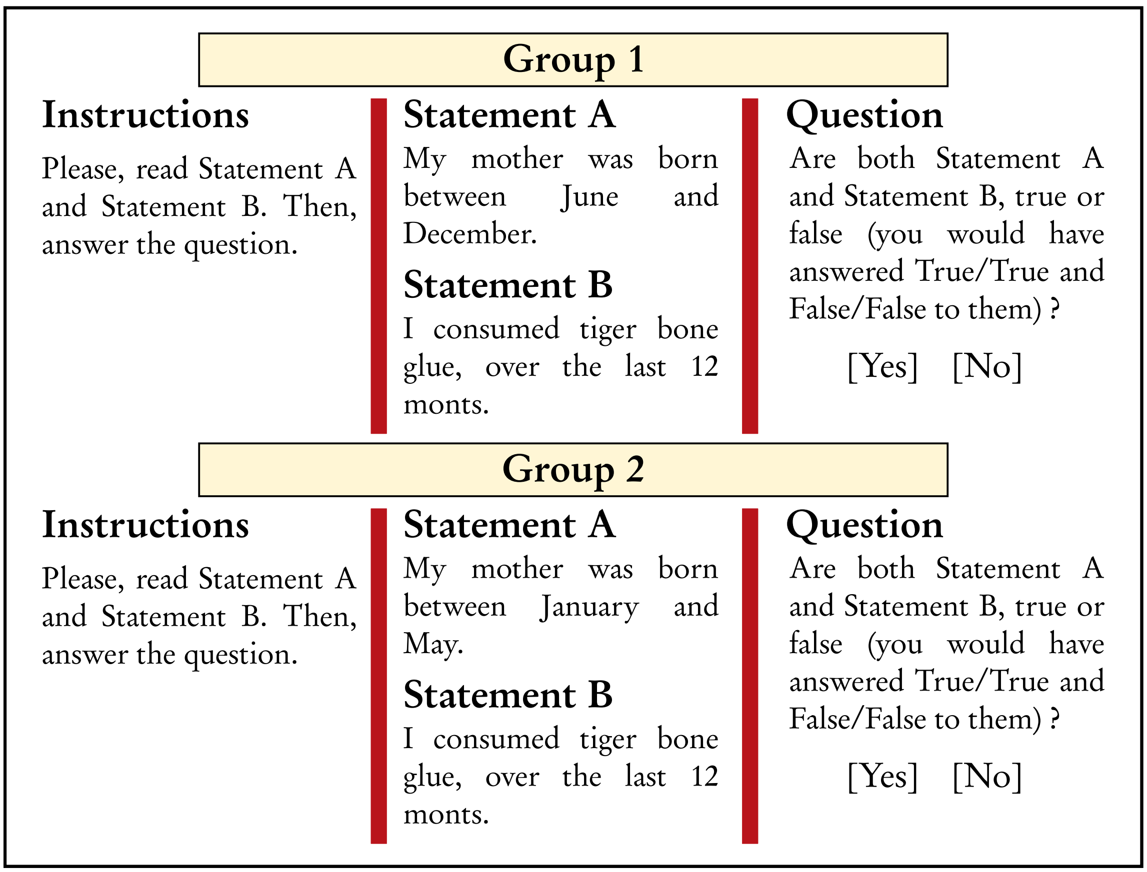

Figure 18 | An example instruction card for the extended crosswise method, with a hypothetical case study about tiger bone glue consumption in Vietnam. Respondents are allocated at random to Group 1 or Group 2. The temporal distribution of births, and therefore the frequency of Statement A is known. Respondents could complete a self-administered questionnaire or answer to an interviewer.

The dual non-randomized response technique and the alternating non-randomized response technique have been proposed to actively account for deception in the triangular model (Wu & Tang, 2016). In the dual non randomized response technique, the sample is split into two groups, with two different non-sensitive questions. The method proceeds exactly as it would normally, but by incorporating the split groups and different sensitive questions, noncompliance can be measured (Fig. 19).

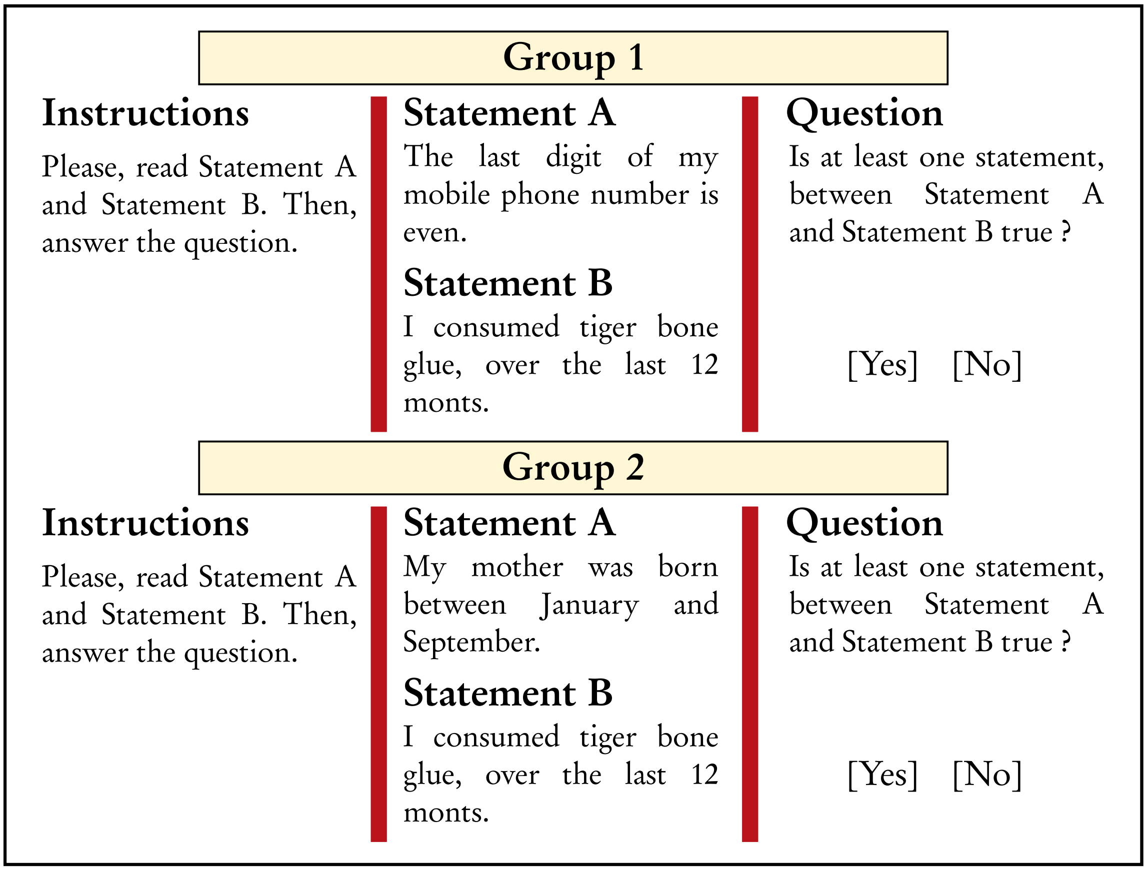

Figure 19 | An example instruction card for the dual non randomized response technique, with a hypothetical case study about tiger bone glue consumption in Vietnam. Respondents are allocated at random to Group 1 or Group 2. The frequency of Statement A is known, but Statement A in the two groups should have a different probability. Respondents could complete a self-administered questionnaire or answer to an interviewer.

In the alternating non randomized response technique, the sample is also split into two groups, but only one non-sensitive question is used (Fig. 20). A key difference from the dual non randomized response technique is that the triangle is "flipped" in one of the groups, in that the respondent marks outside of the triangle if they do not have the sensitive characteristic, but would say yes to the non-sensitive question. Both the two methods have been argued to provide more accurate estimates than the triangular model, with the potential to capture double the "true" prevalence of the behavior in question. However, they still suffer from high variability in prevalence estimation, even for high sample sizes (n $\sim$ 1000; Wu & Tang, 2016).

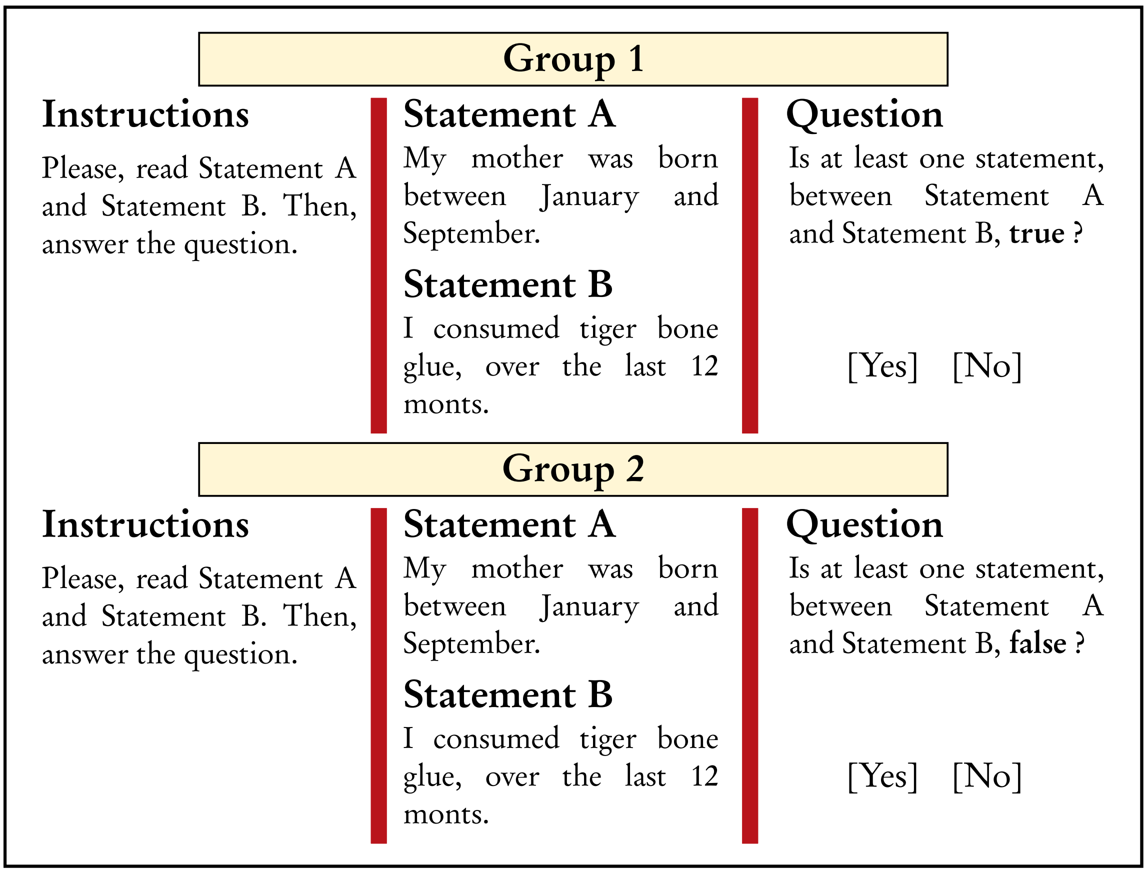

Figure 20 | An example instruction card for the alternating non-randomized response technique (alternative non-randomized response technique), with a hypothetical case study about tiger bone glue consumption in Vietnam. Respondents are allocated at random to Group 1 or Group 2. The frequency of Statement A is known. Respondents could complete a self-administered questionnaire or answer to an interviewer.

3.3 Other SQTs

There are two techniques that were not considered in Nuno and St John (2015), despite having been invented before 2015, and which we note could be adopted to measure sensitive questions in conservation.

3.3.1 Endorsement experiments

Endorsement experiments measure sensitive attitudes. In endorsement experiments, participants are randomly assigned to a treatment and a control group, both expressing their support for a policy on a traditional 5- or 7- points bipolar scale (Blair, Imai, & Lyall 2014; Rosenfeld, Imai, & Shapiro 2016). However, the treatment groups specifically recall a certain issue, which could influence the evaluation of the baseline scenario (Fig. 21).

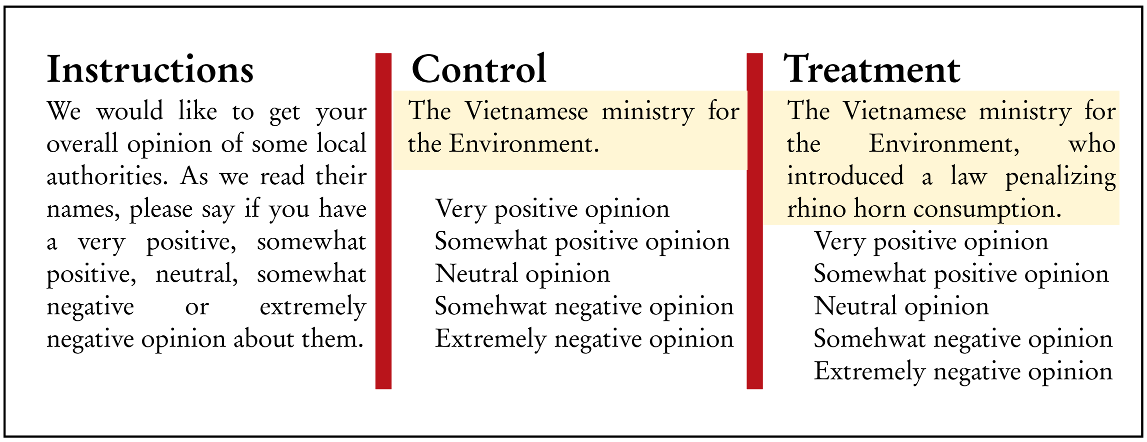

Figure 21 | An example instruction card for an endorsing experiment, from a hypothetical case study about rhino horn consumption in South-East Asia. Respondents are allocated at random to Group 1 or Group 2.

The difference in scores between these two groups reveals the importance of the specific issue at hand. In endorsement experiments individuals never need to reveal their own preferences, and the technique is quite easy to design and administer. However, endorsement experiments are even more statistically inefficient than UCT, which is already characterized by overly high variability (e.g. Davis et al., 2020), and as such large samples are often needed (Hinsley, Keane, St. John, Ibbett, & Nuno 2019; Rosenfeld, Imai, & Shapiro 2016).

3.3.2 The ballot box method

In the ballot box method (Gregson, Zhuwau, Ndlovu, & Nyamukapa 2002) respondents are provided with voting tokens, with a color and a number corresponding to the various questions. An enumerator then explains the technique and reads the questions. Respondents put their tokens into a sealed box, which is opened at the end of the study.

The ballot box method is similar in terms of privacy-protecting mechanism to the bean method (Lau, Yeung, Mui, Tsui, & Gui 2011): respondents provide their anonymous answers and their privacy is protected by the sealed vessel, which is opened at the end of the study (in the case of comparable bean method, the vessel is opened at the end of every day, see Jones, Papworth, Keane, Vickery, & St. John 2020). This mechanism is highly effective for measuring sensitive behaviors at public events, where multiple answers are collected at the same time, blurring individual answers (e.g. during meetings with stakeholders). The ballot box could be extremely easy to implement and understand. Bova, Aswani, Farthing, and Potts (2018), adopted the BB in face-to-face surveys structured interviews with recreational anglers in South Africa, and found that the technique provided higher estimates of non-compliant behaviors than RRT and direct questioning.

3.4 New SQTs

3.4.1 The parallel model

The parallel model (Tian 2014), is a non-randomized technique combining two non-sensitive questions, with known prevalence, with a sensitive question of interest. The parallel model is based on two questions (Fig. 22).

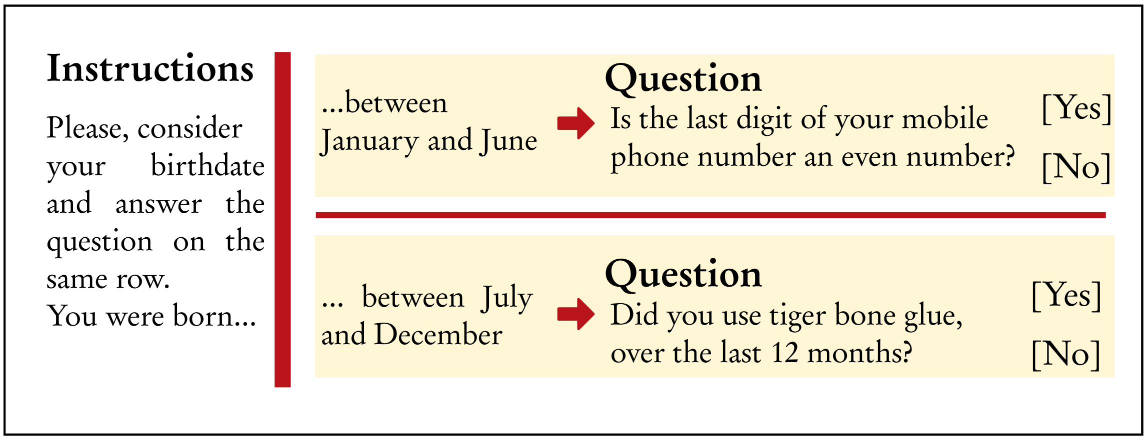

Figure 22 | An example instruction card for the parallel model, containing a hypothetical case study about tiger bone glue usage in Vietnam. The temporal distribution of births and the distribution of mobile phone numbers are known. Respondents could record their answers on a self-administered questionnaire (like in this example) or answer to an interviewer.

In this case, both the distribution of births through the year, and the distribution of cell phone digits, are known and provide information to estimate the prevalence of chimpanzee meat consumption. Tian (2014) claims the parallel model to be statistically more efficient, and more privacy protecting, than the crosswise model and the triangular model. However, to date, the technique has only once been applied in field conditions (e.g. sexual habits, Tian, Liu, & Tang 2019) and there have been no applications of it in conservation.

3.4.2 The "pair" and "list" methods

The pair method and the list method (Lagerås & Lindholm, 2020) apply to multiple choices. In the pair method the respondent is asked to report his/her chosen option and another option, chosen at random, and to write down the two answers in random order (Fig. 23).

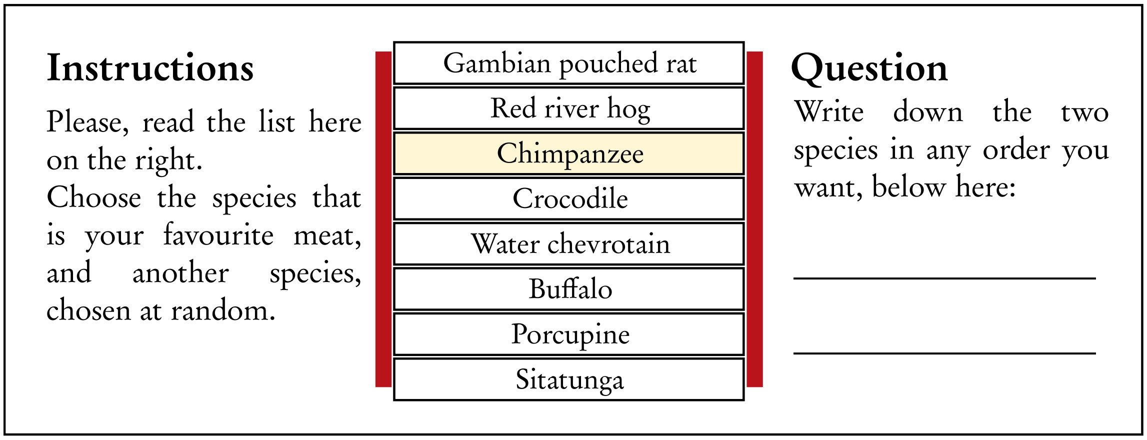

Figure 23 | An example instruction card for the pair method, containing a hypothetical case study about preferences for chimpanzee meat consumption, compared to other bushmeat species (from van Vliet, Nebesse, & Nasi 2015). Respondents could record their answers on a self-administered questionnaire (like in this example) or answer to an interviewer. Alternatively, they can also check the two species on the list, with a cross-mark. To increase privacy protection, lists can be extended to a higher number of species.

In the list method the respondent is presented with a list of options and to indicate whether his/her option is on the list, with different respondents that are shown different lists (Fig. 24). It is easier than the pair method, as it does not require any randomization by the respondent, and privacy protection is even higher. The two methods have never been applied in conservation. The list method is similar to UCT in its random allocation of sets of answers to two groups of respondents. Therefore, like UCT, it might need large sample sizes (n > 1000-2000) to reduce variance (Hinsley, Keane, St. John, Ibbett, & Nuno 2019). The two methods seem suitable to measure preferences when many options are available, as this is their cornerstone of privacy protection, and therefore unsuitable for researchers investigating behaviors with few categories.

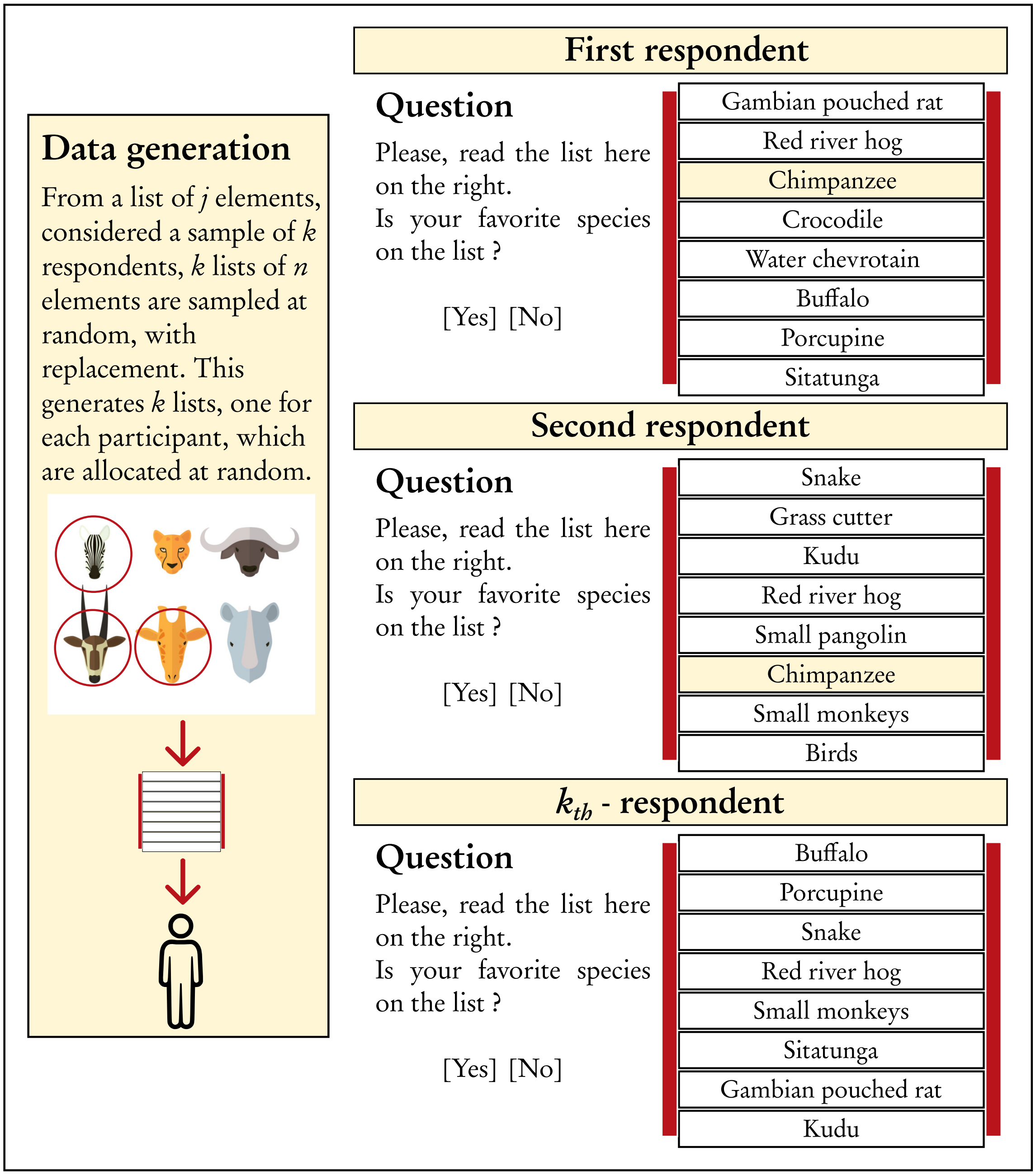

Figure 24 | An example instruction card for the list method, containing a hypothetical case study about preferences for chimpanzee meat consumption, compared to other bushmeat species (van Vliet, Nebesse, & Nasi 2015). Respondents could record their answers on a self-administered questionnaire (like in this example) or answer to an interviewer. Each respondent has a different list, generated at random. Lists might contain sensitive questions, or not.

3.4.3 The person count technique

In the person count technique (Wolter, 2019), respondents are randomly assigned to a control and a treatment group. In the control group, respondents are asked to think of three people that they know, and state how many of those three would agree with a sensitive statement (e.g. "*I use pangolin scales for my health*"). In the treatment group, respondents are asked to think of the same three people from their social group, but should also consider themselves and their own preferences when stating the number of people who agree with the sensitive statement. The person count technique overcomes an issue of the nominative technique which is that estimations of how many individuals in a social group perform a sensitive act is challenging to identify, can be inflated and/or vague, and don't concretely capture the actual preferences of the individual being interviewed (e.g. Davis et al., 2019). However, the person count technique could be highly sensitive to floor and ceiling effects, and due to the separation of the sample into two groups, it could suffer from large standard errors, like UCT (Wolter, 2019).

4 Discussion

Worldwide, there is a growing need to quantify and understand conservation-negative behaviors and potential noncompliance with conservation regulations, so as to better address and reduce unsustainable resource use (Solomon, Gavin, & Gore 2015). SQTs are especially attractive to conservation social scientists and conservationists seeking to use social science, as they promise to measure human behavior and attitudes in an unbiased way. Our findings indicate that SQTs have blossomed over the last five years, with both the invention of entirely new methods and the development of many new applications of existing techniques. It is encouraging to see that this development has gone beyond RRT, the most common SQTs method, and involves methods which do not require any randomization device. This in particular is promising for being more easily adopted by conservationists working in the field. Some methods, like the ballot box method and bean method, while not having a sophisticated protection mechanism, do keep responses private and appear to be simple and adaptable to many situations (Bova, Aswani, Farthing, & Potts, 2018; Jones, Papworth, Keane, Vickery, & St. John 2020).

SQTs still pose some major challenges to researchers. First, while they can eliminate the risk of privacy disclosure and sanctions, they do not entirely resolve respondents' concerns about the misinterpretation of their answers, or when answering about taboos (e.g. St. John et al., 2018). Second, due to their cognitive load, and their sometimes complex functioning, SQTs can in some situations not be understood by respondents (e.g. Davis et al., 2019). Third, which method is "most" trusted seems to vary across different contexts, with RRT found to be "untrustworthy" in some studies (Hoffmann, De Puiseau, Schmidt, & Musch 2017), and a more trusted method in others (Davis et al., 2019). Fourth, other studies have suggested that the interpretation of SQTs might be tricky even for highly educated respondents (Jerke, Johann, Rauhut, & Thomas 2019).

In recognition of these uncertainties, we call for a push, particularly in conservation, of greater experimentation with SQTs in the field. Many of the new methods and variants identified in this study have received relatively little or no practical application. They have been deemed statistically valid; however, they may not in fact be scientifically valid, if they are created based on erroneous assumptions about how humans act (Navarro, 2019). Testing these methods with field experiments, coupled with qualitative research around the sensitive behavior of interest (Chuang, Dupas, Huillery, & Seban 2019) is ultimately the sole way to diagnose potential problems and to test for their performances. For example, Davis, Crudge, and Glikman (2020) found that nominative technique was entirely unnecessary in a study on bear bile use in northern Laos, due to a complete lack of sensitivity around use. Most of the non-randomized techniques reported in Table 1 have received little application in the field, and considered their potential for conservation, their testing should be a priority.

Another relevant limitation for many SQTs is sample size. Most techniques require thousands of observations, and even more for cheating detection (Clark & Desharnais, 1998; Schröter et al., 2016), yet still they can suffer from large estimation errors. Even many recent techniques achieve statistical efficiency only at the price of their practical implementation, through requiring large sample sizes that ultimately become cumbersome (e.g. multinomial RRT, double list item sum technique, pair and list method). Especially for conservationists, large samples are often beyond what can be achieved, due to money and time complications in gaining large samples. Moreover, large samples and standard errors could limit the application of SQTs in longitudinal designs. To understand whether small sample sizes could perhaps be effective for some of these techniques, there is a growing need for practical experimentation and subsequently clear guidelines about sample size requirements of the various SQTs. For instance, at least two studies using RRT have shown that although large error sizes persist, samples that are considerably lower than 1000 can still result in sensible estimations of a behavior ($\sim$ 500 individuals: Ruppert et al., 2020; $\sim$ 200 individuals: St. John, Mai, & Pei 2015). Evidence-based guidelines around such a fundamental aspect of studies would ultimately assist researchers to decide whether, and which, SQTs are feasible for their study.

Moreover, we encourage conservationists to carefully think about why they need SQTs. While many studies aim to quantify the prevalence of sensitive behaviors, many others focus on understanding their drivers. These usually model the influence of covariates over the probability that respondents engage in a certain behavior, but SQTs might not be the best tool for this scope. Many of them do not allow for covariates at all (e.g. bean method, ballot box method, nominative technique, person count method) and even those that can generally further worsen their statistical power when doing so. Moreover, while there are implementations of the RRT for randomizing questions to be used as a response variable or as a predictor in a regression model (Cruyff, Böckenholt, Van Der Heijden, & Frank 2016), there is no way of randomizing both. There are limits to adding noise.

We believe that other methods can be suitable for measuring drivers of sensitive, or deviant, behaviors. For example endorsement experiments, but also choice experiments or factorial survey experiments (Auspurg & Hinz, 2014), can be valuable approaches for understanding drivers of individual behavior in conservation, even for sensitive topics.

5 Conclusions

There are a wealth of SQTs that conservationists can use to monitor and understand sensitive topics. Although there are challenges intrinsic in many, if not most, of the methods described here, some may be more applicable to researchers than classic forms of UCT and RRT. We urge researchers to explore other options for reducing biases, but caution that researchers should ideally have a fundamental understanding of the sensitive behavior and the system it's in, before plunging into use of an SQT. Moreover, we caution overt reliance on complex statistical techniques, particularly models, without adequate interrogation of the assumptions and structure (e.g. MacNally, 2000). Finally, where possible, we suggest that simple measures of assessing internal consistency should be applied (Chuang, Dupas, Huillery, & Seban 2019).

Although there is no perfect SQT, we believe that understanding sensitive behaviors is highly important. Despite the hurdles intrinsic in obtaining an accurate estimation of behavior, we believe that the diversity of SQTs indicate that the use of them should be achievable for any researcher in conservation, regardless of sample size or funds. The use of them must be careful and thoughtful; however, the resulting reward of a more complete understanding of the complex and obscure counterbalance the effort involved in ensuring SQT efficacy.

Acknowledgements

With thanks to E. Moran for support in writing this paper.

Table 1. Groups of SQTs that we identified, classified on the basis of their privacy-protection mechanism. Advances in existing techniques, as well as new SQTs are indicated as "new!".

Privacy protection	Techniques	Outcomes	Covariates
Noise with a known distribution is added to individual answers, usually through a randomizing device (e.g. a die, a deck of cards). Privacy is protected by adding noise at the individual level, respondents should feel protected and provide honest answers.	Forced response RRT (Boruch, 1971) • Multidimensional RRT (Cruyff, Böckenholt, & van der Heijden 2016) **new!** • Quantitative RRT (Cao, Breidt, Solomon, Conteh, & Gavin 2018) **new**! Unrelated question RRT (Greenberg, 1969) • Optional unrelated question (Chhabra, Dass, & Gupta 2016) **new!** • Cheater detection model (Schröter et al. 2016) **new!** • RRT-crosswise design (Tu & Hsieh, 2017) **new**! Mirrored design RRT (Warner, 1965) Disguised response RRT (Kuk, 1990)	Binary Ordinal scales Discrete counts Multinomial scales	Yes
Respondents are asked to provide an overall answer about a set of behaviors, including sensitive and non-sensitive ones. They might indicate how many of these behaviors apply to them (UCT), or summing the frequencies of individual answers (IST). The difference between groups estimates the frequency of the behavior of interest. Privacy is protected by the fact that respondents provide an overall answer about a set of 4 - 10 behaviors.	Unmatched count technique (UCT, Droitcour et al., 2004) • Double list UCT (Glynn, 2013) • Single-sample count UCT (Petróczi et al., 2011) • Longitudinal UCT (Gaia & Al Baghal, 2019) **new!** • Negative binomial/Poisson UCT (Tian, Tang, Wu, & Liu 2017) **new!** • Poisson-Poisson UCT (Liu, Tian, Wu, & Tang 2019) **new!** Item sum technique (IST, Trappmann, Krumpal, Kirchner, & Jann 2014) • Double list IST (Krumpal, Jann, Korndörfer, & Schmukle 2018) **new!**	Binary (UCT) Discrete counts (IST)	Yes

Privacy protection	Techniques	Outcomes	Covariates
Respondents provide a conjoint answers to two, or more, questions. These include the sensitive question of interest and one, or more, non-sensitive questions whose prevalence in the population is known. Privacy is protected by the fact that respondents provide a conjoint answer (e.g. in crosswise model they indicate if they provided two different answers to the two questions, or not) and one of the two questions is sensitive.	Crosswise model (Yu, Tian, & Tang 2008) • Extended Crosswise Model (Heck, Hoffmann, & Moshagen 2018) **new!** • RRT-crosswise design (Tu & Hsieh 2017) **new!** Triangular model (Yu, Tian, & Tang 2008) • Dual non-randomized response (Wu & Tang 2016) **new!** • Alternating non-randomized response (Wu & Tang 2016) **new!** Diagonal model (Groenitz, 2014) Hidden sensitivity model (Tian, Yu, Tang, & Geng 2007) Parallel model (Tian, 2014) **new!**	Binary Discrete counts	Yes
Respondents are allocated to two groups, each one containing three blocks. The two groups contain a list with the one sensitive and various non-sensitive behaviors, which are assigned at random to the three blocks, so that Group1 and Group2 have the same overall behaviors, but coupled differently. Respondents indicate only the block (A, B or C) to which they belong, without indicating their real behavior. Privacy is protected because researchers do not know which is the behavior practiced by respondents.	Grouped-answer method (Droitcur & Larson 2002)	Binary	No
Respondents are asked questions posed in a negative direction (e.g. *'I did not fish ...'*), so they do not answer any sensitive question which could jeopardize them.	Negative question (Esponda & Guerrero 2009)	Binary Ordinal Multinomial	Yes

Privacy protection	Techniques	Outcomes	Covariates
Respondents are provided with two jars of multicolored beans. They are asked to move a black bean from the small to the large jar, if the answer to a question is “no” and a bean of a certain color, if the answer to a question is “yes”. After many answers, beans are counted. Privacy is protected because (*i*) jars have many beans, so that the movement of a single bean goes unnoticed, (*ii*) jars are opened at the end of the study and (*iii*) respondents are not observed by researchers, when they move beans.	Bean method (Lau, Yeung, Mui, Tsui, & Gu 2011)	Binary	No
Respondents are allocated to two groups, each one containing three blocks. The two groups contain a list with the one sensitive and various non-sensitive behaviors, which are assigned at random to the three blocks, so that Group1 and Group2 have the same overall behaviors, but coupled differently. Respondents indicate only the block (A, B or C) to which they belong, without indicating their real behavior. Privacy is protected because researchers do not know which is the behavior practiced by respondents.	Grouped-answer method (Droitcur & Larson 2002)	Binary	No

Privacy protection	Techniques	Outcomes	Covariates
Pair method : respondents answer the sensitive question and a safe question, picked from a list. The two answers are then written down in a random order and privacy is protected by private randomization made by respondents. List method: respondent is presented with a list of options and indicate if the sensitive option is on the list. Each respondent receives a different list of options, extracted at random..	List method (Lagerås & Lindholm 2020) **new!** Pair method (Lagerås & Lindholm 2020) **new!**	Multinomial	No
Respondents indicate how many people in their social network engage into a sensitive behavior. Then, one of reported people is picked at random and are asked how many people they believe to know about their friend's behavior (nominative technique). Alternatively, they are split in a control and a treatment group (person count technique). In the treatment group, respondents are asked to add themselves to the number of indicated people. Privacy is protected because researchers do not know the real number of people nominated by respondents.	Nominative technique (Miller, 1985) Person count technique (Wolter, 2019) **new!**	Discrete	No

Privacy protection	Techniques	Outcomes	Covariates
Respondents are provided with voting tokens, with a color and a number corresponding to the answers to the various questions. An enumerator then explains the technique and reads the questions. Respondents put their tokens sheet into a sealed box, which is opened at the end of the study. Privacy is protected because the urn is opened well after the interview and it is impossible to trace back individual answers.	Ballot box method (Gregson, Zhuwau, Ndlovu, & Nyamukapa 2002)	Binary Ordered Discrete	No
Respondents are randomly allocated to a control and a treatment condition. The control condition illustrates a certain issue (e.g. a policy), while treatment condition highlights a certain aspect of interest of that issue (e.g. policy impacts over wildlife trade). Respondents express their approval of the subject described, ignoring respondents in the other condition. The difference in scores between the two groups reflects the extent to which a certain issue is endorsed by respondents. Privacy is not protected, but framing might make respondents' unaware of questions' sensitivity, as a certain question could appear sensitive only if compared to other questions in a survey.	Endorsement experiments (Blair, Imai, & Lyall 2014)	Ordered (usually bipolar)	No

References

- Arnab, R., & Rueda, M. (2016). Optional randomized response: A critical review. Handbook of Statistics, 34, 253-271.https://doi.org/10.1016/bs.host.2016.01.014
- Auspurg, K., & Hinz, T. (2014). Factorial survey experiments (1st ed.). Thousand Oaks, CA: Sage Publications.
- Bennett, N. J., Roth, R., Klain, S. C., Chan, K., Christie, P., Clark, D. A., Cullman, G., Curran, D., Durbin, T. J., Epstein, G., Greenberg, A., Nelso, M. P., Sandlons, J., Stedman, R., Teel, T. L., Thomas, R., Veríssimo, D., & Wyborn, C. (2017). Conservation social science: Understanding and integrating human dimensions to improve conservation. Biological conservation, 205, 93-108. https://doi.org/10.1016/j.biocon.2016.10.006
- Blair, G., Imai, K., & Lyall, J. (2014). Comparing and combining list and endorsement experiments: Evidence from Afghanistan. American Journal of Political Science, 58(4), 1043-1063. https://doi.org/10.1111/ajps.12086
- Blair, G., Imai, K., & Zhou, Y. Y. (2015). Design and analysis of the randomized response technique. Journal of the American Statistical Association, 110(511), 1304-1319. https://doi.org/10.1080/01621459.2015.1050028
- Blair, G., & Imai, K. (2012). Statistical analysis of list experiments. Political Analysis, 20(1), 47-77. https://doi.org/10.1093/pan/mpr048
- Bornmann, L., & Mutz, R. (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. Journal of the Association for Information Science and Technology, 66(11), 2215-2222. https://doi.org/10.1002/asi.23329
- Boruch, R. F. (1971). Assuring confidentiality of responses in social research: A note on strategies. The American Sociologist, 6, 308-311.
- Bova, C. S., Aswani, S., Farthing, M. W., & Potts, W. M. (2018). Limitations of the random response technique and a call to implement the ballot box method for estimating recreational angler compliance using surveys. Fisheries Research, 208, 34-41. https://doi.org/10.1016/j.fishres.2018.06.017
- Brittain, S., Ibbett, H., de Lange, E., Dorward, L., Hoyte, S., Marino, A., Milner-Gulland, J., Newth, J., Rakotonarivo, S., Veríssimo, D., & Lewis, J. (2020). Ethical considerations when conservation research involves people. Conservation Biology, 34(4), 925-933. https://doi.org/10.1111/cobi.13464
- Camilotti, G. (2016). Interventions to stop female genital cutting and the evolution of the custom: Evidence on age at cutting in Senegal. Journal of African Economies, 25(1), 133-158. https://doi.org/10.1093/jae/ejv013
- Cao, M., Breidt, F. J., Solomon, J. N., Conteh, A., & Gavin, M. C. (2018). Understanding the drivers of sensitive behavior using Poisson regression from quantitative randomized response technique data. PloS one, 13(9), e0204433. https://doi.org/10.1371/journal.pone.0204433
- Cerri, J., Ciappelli, A., Lenuzza, A., Nocita, A., & Zaccaroni, M. (2018). The randomised response technique: A valuable approach to monitor pathways of aquatic biological invasions. Fisheries Management and Ecology, 24(6), 504-511. https://doi.org/10.1111/fme.12258
- Cerri, J., Mori, E., Vivarelli, M., & Zaccaroni, M. (2017). Are wildlife value orientations useful tools to explain tolerance and illegal killing of wildlife by farmers in response to crop damage?. European Journal of Wildlife Research, 63(4), 70. https://doi.org/10.1007/s10344-017-1127-0
- Cerri, J., Scuffi, L., Nocita, A., Zaccaroni, M., Lenuzza, A., & Cruyff, M. (2018). Characterizing noncompliance in conservation: a multidimensional Randomized Response Technique for multinomial responses. BioRxiv, 453159. https://doi.org/10.1101/453159
- Chhabra, A., Dass, B. K., & Gupta, S. (2016). Estimating prevalence of sexual abuse by an acquaintance with an optional unrelated question RRT model. The North Carolina Journal of Mathematics and Statistics, 2, 1-9.
- Chang, C. H., Cruyff, M. J., & Giam, X. (2018). Examining conservation compliance with randomized response technique analyses. Conservation Biology, 32(6), 1448-1456. https://doi.org/10.1111/cobi.13133
- Chuang, E., Dupas, P., Huillery, E., & Seban, J. (2019). Sex, lies, and measurement: Do indirect response survey methods work. Working Paper. https://web.stanford.edu/ pdupas/CDHS_measurement.pdf
- Clark, S. J., & Desharnais, R. A. (1998). Honest answers to embarrassing questions: Detecting cheating in the randomized response model. Psychological Methods, 3(2), 160. https://psycnet.apa.org/doi/10.1037/1082-989X.3.2.160
- Cialdini, R. B., Demaine, L. J., Sagarin, B. J., Barrett, D. W., Rhoads, K., & Winter, P. L. (2006). Managing social norms for persuasive impact. Social influence, 1(1), 3-15. https://doi.org/10.1080/15534510500181459
- Cruyff, M. J., Böckenholt, U., & van der Heijden, P. G. (2016). The multidimensional randomized response design: Estimating different aspects of the same sensitive behavior. Behavior research methods, 48(1), 390-399. https://doi.org/10.3758/s13428-015-0583-2

- Cruyff, M. J., Böckenholt, U., Van Der Heijden, P. G., & Frank, L. E. (2016). A review of regression procedures for randomized response data, including univariate and multivariate logistic regression, the proportional odds model and item response model, and self-protective responses. Handbook of Statistics, 34, 287-315. https://doi.org/10.1016/bs.host.2016.01.016
- Davis, E. O., Crudge, B., Lim, T., O'Connor, D., Roth, V., Hunt, M., & Glikman, J. A. (2019). Understanding the prevalence of bear part consumption in Cambodia: A comparison of specialised questioning techniques. PLoS One, 14(2), e0211544. https://doi.org/10.1371/journal.pone.0211544
- Davis, E. O., & Glikman, J. A. (2020). An assessment of wildlife use by Northern Laos nationals. Animals, 10(4), 685. https://doi.org/10.3390/ani10040685
- Davis, E.O., Willemsen, M., Dang, V., O'Connor, D., & Glikman, J.A. (2020). An updated analysis of the consumption of tiger products in urban Vietnam. Global Ecology and Conservation, 22, p.e00960. https://doi.org/10.1016/j.gecco.2020.e00960
- Davis, E.O., Veríssimo, D., Crudge, B., Lim, T., Roth, V., & Glikman, J.A. (2020). Insights for reducing the consumption of wildlife: the use of bear bile and gallbladder in Cambodia. SocArxiv. https://osf.io/preprints/socarxiv/kpw27
- Davis, E.O., Crudge, B., & Glikman, J.A. (2020). Nominative technique, a simple tool for assessing illegal wildlife consumption. SocArxiv. https://osf.io/preprints/socarxiv/gvknx
- Diekmann, A. (2012). Making use of "Benford's Law" for the randomized response technique. Sociological Methods & Research, 41(2), 325-334. https://doi.org/10.1177%2F0049124112452525
- Droitcour, J., Caspar, R. A., Hubbard, M. L., Parsley, T. L., Visscher, W., & Ezzati, T. M. (2004). The item count technique as a method of indirect questioning: A review of its development and a case study application. Measurement errors in surveys, 185-210. https://doi.org/10.1002/9781118150382.ch11
- Droitcour, J. A., & Larson, E. M. (2002). An innovative technique for asking sensitive questions: The three-card method. Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, 75(1), 5-23. https://doi.org/10.1177%2F075910630207500103
- Eichhorn, B. H., & Hayre, L. S. (1983). Scrambled randomized response methods for obtaining sensitive quantitative data. Journal of Statistical Planning and inference, 7(4), 307-316. https://doi.org/10.1016/0378-3758(83)90002-2
- Esponda, F., & Guerrero, V. M. (2009). Surveys with negative questions for sensitive items. Statistics & Probability Letters, 79(24), 2456-2461. https://doi.org/10.1016/j.spl.2009.08.019
- Gaia, A., & Al Baghal, T. (2019). The Longitudinal Item Count Technique: a new technique for asking sensitive questions in surveys. Methods, data, analyses: a journal for quantitative methods and survey methodology (mda), 13(1), 111-137. https://doi.org/10.12758/mda.2018.09
- Gino, F., & Ariely, D. (2016). Dishonesty explained: What leads moral people to act immorally. The social psychology of good and evil, 322-344. http://people.duke.edu/ dandan/webfiles/PapersDisHonesty/Dishonesty%20Explained%20What.doc
- Glynn, A. N. (2013). What can we learn with statistical truth serum? Design and analysis of the list experiment. Public Opinion Quarterly, 77(S1), 159-172. https://doi.org/10.1093/poq/nfs070c
- Greenberg, B. G., Abul-Ela, A. L. A., Simmons, W. R., & Horvitz, D. G. (1969). The unrelated question randomized response model: Theoretical framework. Journal of the American Statistical Association, 64(326), 520-539. https://doi.org/10.1080/01621459.1969.10500991
- Gregson, S., Zhuwau, T., Ndlovu, J., & Nyamukapa, C. A. (2002). Methods to reduce social desirability bias in sex surveys in low-development settings: experience in Zimbabwe. Sexually transmitted diseases, 29(10), 568-575.
- Groenitz, H. (2014). A new privacy-protecting survey design for multichotomous sensitive variables. Metrika, 77(2), 211-224. https://doi.org/10.1007/s00184-012-0406-8
- Heck, D. W., Hoffmann, A., & Moshagen, M. (2018). Detecting nonadherence without loss in efficiency: A simple extension of the crosswise model. Behavior research methods, 50(5), 1895-1905. https://doi.org/10.3758/s13428-017-0957-8
- Hinsley, A., Keane, A., St. John, F. A., Ibbett, H., & Nuno, A. (2019). Asking sensitive questions using the unmatched count technique: Applications and guidelines for conservation. Methods in Ecology and Evolution, 10(3), 308-319. https://doi.org/10.1111/2041-210X.13137

- Hinsley, A., Nuno, A., Ridout, M., St. John, F. A., & Roberts, D. L. (2017). Estimating the extent of CITES noncompliance among traders and end-consumers; lessons from the global orchid trade. Conservation Letters, 10(5), 602-609. https://doi.org/10.1111/conl.12316
- Hoffmann, A., De Puiseau, B. W., Schmidt, A. F., & Musch, J. (2017). On the comprehensibility and perceived privacy protection of indirect questioning techniques. Behavior Research Methods, 49(4), 1470-1483. https://doi.org/10.3758/s13428-016-0804-3
- Hoffmann, A., Diedenhofen, B., Verschuere, B., & Musch, J. (2015). A strong validation of the crosswise model using experimentally-induced cheating behavior. Experimental Psychology. https://doi.org/10.1027/1618-3169/a000304
- Hoffmann, A., Meisters, J., & Musch, J. (2020). On the validity of non-randomized response techniques: an experimental comparison of the crosswise model and the triangular model. Behavior Research Methods, 1-15. https://doi.org/10.3758/s13428-020-01349-9
- Höglinger, M., & Diekmann, A. (2017). Uncovering a blind spot in sensitive question research: false positives undermine the crosswise-model RRT. Political Analysis, 25(1), 131-137. https://doi.org/10.1017/pan.2016.5
- Ibbett, H., Lay, C., Phlai, P., Song, D., Hong, C., Mahood, S. P., & Milner-Gulland, E. J. (2019). Conserving a globally threatened species in a semi-natural, agrarian landscape. Oryx, 53(1), 181-191. https://doi.org/10.1017/S0030605316001708
- Jerke, J., Johann, D., Rauhut, H., & Thomas, K. (2019). Too sophisticated even for highly educated survey respondents? A qualitative assessment of indirect question formats for sensitive questions. Survey Research Methods, 13(3), 319-351. https://doi.org/10.18148/srm/2019.v13i3.7453
- Jones, S., Papworth, S., Keane, A. M., Vickery, J., & St John, F. A. (2020). The bean method as a tool to measure sensitive behaviour. Conservation Biology. https://doi.org/10.1111/cobi.13607
- Jones, J. P., Andriamarovololona, M. M., & Hockley, N. (2008). The importance of taboos and social norms to conservation in Madagascar. Conservation biology, 22(4), 976-986. https://doi.org/10.1111/j.1523-1739.2008.00970.x
- Korndörfer, M., Krumpal, I., & Schmukle, S. C. (2014). Measuring and explaining tax evasion: Improving self-reports using the crosswise model. Journal of Economic Psychology, 45, 18-32. https://doi.org/10.1016/j.joep.2014.08.001
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. Quality & Quantity, 47(4), 2025-2047. https://doi.org/10.1007/s11135-011-9640-9
- Krumpal, I., Jann, B., Korndörfer, M., & Schmukle, S. (2018). Item sum double-list technique: An enhanced design for asking quantitative sensitive questions. Survey Research Methods, 12(2), 91-102. https://doi.org/10.18148/srm/2018.v12i2.7247
- Kuk, A. Y. (1990). Asking sensitive questions indirectly. Biometrika, 77(2), 436-438. https://doi.org/10.1093/biomet/77.2.436
- Lagerås, A., & Lindholm, M. (2020). How to ask sensitive multiple-choice questions. Scandinavian Journal of Statistics, 47(2), 397-424. https://doi.org/10.1111/sjos.12411
- Lau, J. T., Yeung, N. C., Mui, L. W., Tsui, H. Y., & Gu, J. (2011). A simple new method to triangulate self-reported risk behavior data—the bean method. Sexually transmitted diseases, 38(9), 788-792. https://doi.org/10.1097/OLQ.0b013e318218cc66
- Lee, S. M., Peng, T. C., Tapsoba, J. D. D., & Hsieh, S. H. (2017). Improved estimation methods for unrelated question randomized response techniques. Communications in Statistics-Theory and Methods, 46(16), 8101-8112. https://doi.org/10.1080/03610926.2016.1175626
- Liu, Y., Tian, G. L., Wu, Q., & Tang, M. L. (2019). Poisson–Poisson item count techniques for surveys with sensitive discrete quantitative data. Statistical Papers, 60(5), 1763-1791. https://doi.org/10.1007/s00362-017-0895-7
- MacNally, R. (2000). Regression and model-building in conservation biology, biogeography and ecology: the distinction between–and reconciliation of–'predictive'and 'explanatory'models. Biodiversity & Conservation, 9(5), 655-671. https://doi.org/10.1023/A:1008985925162
- Meisters, J., Hoffmann, A., & Musch, J. (2020). Can detailed instructions and comprehension checks increase the validity of crosswise model estimates?. PloS one, 15(6), e0235403. https://doi.org/10.1371/journal.pone.0235403

- Miller, J. D. (1985). The nominative technique: A new method of estimating heroin prevalence. NIDA Research Monograph, 54, 104-124.
- Navarro, D. J. (2019). Between the devil and the deep blue sea: Tensions between scientific judgement and statistical model selection. Computational Brain & Behavior, 2(1), 28-34. https://doi.org/10.1007/s42113-018-0019-z
- Nuno, A., Bunnefeld, N., Naiman, L. C., & Milner-Gulland, E. J. (2013). A novel approach to assessing the prevalence and drivers of illegal bushmeat hunting in the Serengeti. Conservation Biology, 27(6), 1355-1365. https://doi.org/10.1111/cobi.12124
- Nuno, A., & St John, F. A. (2015). How to ask sensitive questions in conservation: A review of specialized questioning techniques. Biological Conservation, 189, 5-15. https://doi.org/10.1016/j.biocon.2014.09.047
- Perri, P. F., Rodríguez, B. C., & García, M. D. M. R. (2018). A mixed-mode sensitive research on cannabis use and sexual addiction: improving self-reporting by means of indirect questioning techniques. Quality & Quantity, 52(4), 1593-1611. https://doi.org/10.1007/s11135-017-0537-0
- Petróczi, A., Nepusz, T., Cross, P., Taft, H., Shah, S., Deshmukh, N., Schaffer, J., Shane, M., Adesanwo, C., Barker, J., & Naughton, D. P. (2011). New non-randomised model to assess the prevalence of discriminating behaviour: a pilot study on mephedrone. Substance abuse treatment, prevention, and policy, 6(1), 20. https://doi.org/10.1186/1747-597X-6-20
- Pollock, K. H., & Bek, Y. (1976). A comparison of three randomized response models for quantitative data. Journal of the American Statistical Association, 71(356), 884-886.
  https://doi.org/10.1080/01621459.1976.10480963
- Raghavarao, D., & Federer, W. T. (1979). Block total response as an alternative to the randomized response method in surveys. Journal of the Royal Statistical Society: Series B (Methodological), 41(1), 40-45. https://doi.org/10.1111/j.2517-6161.1979.tb01055.x
- Rosenfeld, B., Imai, K., & Shapiro, J. N. (2016). An empirical validation study of popular survey methodologies for sensitive questions. American Journal of Political Science, 60(3), 783-802.
  https://doi.org/10.1111/ajps.12205
- Ruppert, K. A., Sponarski, C. C., Davis, E., Masiaine, S., Larpei, L., Lekalgitele, J., Lekupanai, R., Lekushan, J., Lemirishan, J., Lenaipa, D., Lenyakopiro, J., Lerapayo, S., Lororua, M., Stacy-Dawes, J., & Glikman, J. A. (2020). Use of Specialized Questioning Techniques to Detect Decline in Giraffe Meat Consumption. https://doi.org/10.31235/osf.io/8vw7c
- Russo, A., & Strazzari, F. (2020). The politics of safe research in violent and illiberal contexts. Bliesemann de Guevara, B., & Bøås, M. (Eds.), Doing Fieldwork in Areas of International Intervention: A Guide to Research in Violent and Closed Contexts (pp. 75-94). 75. Bristol, UK: Bristol University Press.
- Santangeli, A., Arkumarev, V., Rust, N., & Girardello, M. (2016). Understanding, quantifying and mapping the use of poison by commercial farmers in Namibia–implications for scavengers' conservation and ecosystem health. Biological Conservation, 204, 205-211. https://doi.org/10.1016/j.biocon.2016.10.018
- Schröter, H., Studzinski, B., Dietz, P., Ulrich, R., Striegel, H., & Simon, P. (2016). A comparison of the Cheater Detection and the Unrelated Question Models: A randomized response survey on physical and cognitive doping in recreational triathletes. PLoS One, 11(5), e0155765. https://doi.org/10.1371/journal.pone.0155765
- Solomon, J., Jacobson, S. K., Wald, K. D., & Gavin, M. (2007). Estimating illegal resource use at a Ugandan park with the randomized response technique. Human Dimensions of Wildlife, 12(2), 75-88. https://doi.org/10.1080/10871200701195365
- Solomon, J. N., Gavin, M. C., & Gore, M. L. (2015). Detecting and understanding non-compliance with conservation rules. Biological Conservation, 189, 1-4. https://doi.org/10.1016/j.biocon.2015.04.028
- St John, F. A., Edwards-Jones, G., Gibbons, J. M., & Jones, J. P. (2010). Testing novel methods for assessing rule breaking in conservation. Biological Conservation, 143(4), 1025-1030.
  https://doi.org/10.1016/j.biocon.2010.01.018
- St. John, F. A, Mai, C. H., & Pei, K. J. C. (2015). Evaluating deterrents of illegal behaviour in conservation: carnivore killing in rural Taiwan. Biological Conservation, 189, 86-94. https://doi.org/10.1016/j.biocon.2014.08.019
- St. John, F. A., Linkie, M., Martyr, D. J., Milliyanawati, B., McKay, J. E., Mangunjaya, F. M., Leader-Williams, N., & Struebig, M. J. (2018). Intention to kill: Tolerance and illegal persecution of Sumatran tigers and sympatric species. Conservation Letters, 11(4), e12451. https://doi.org/10.1111/conl.12451

- St John, F. A., Mai, C. H., & Pei, K. J. C. (2015). Evaluating deterrents of illegal behaviour in conservation: carnivore killing in rural Taiwan. Biological Conservation, 189, 86-94. https://doi.org/10.1016/j.biocon.2014.08.019
- Tian, G. L. (2014). A new non-randomized response model: The parallel model. Statistica Neerlandica, 68(4), 293-323. https://doi.org/10.1111/stan.12034
- Tian, G. L., Liu, Y., & Tang, M. L. (2019). Logistic regression analysis of non-randomized response data collected by the parallel model in sensitive surveys. Australian & New Zealand Journal of Statistics, 61(2), 134-151. https://doi.org/10.1111/anzs.12258
- Tian, G. L., Tang, M. L., Wu, Q., & Liu, Y. (2017). Poisson and negative binomial item count techniques for surveys with sensitive question. Statistical methods in medical research, 26(2), 931-947. https://doi.org/10.1177%2F0962280214563345
- Tian, G. L., Yu, J. W., Tang, M. L., & Geng, Z. (2007). A new non-randomized model for analysing sensitive questions with binary outcomes. Statistics in Medicine, 26(23), 4238-4252. https://doi.org/10.1002/sim.2863
- Trappmann, M., Krumpal, I., Kirchner, A., & Jann, B. (2014). Item sum: a new technique for asking quantitative sensitive questions. Journal of Survey Statistics and Methodology, 2(1), 58-77. https://doi.org/10.1093/jssam/smt019
- Tu, S. H., & Hsieh, S. H. (2017). Estimates of lifetime extradyadic sex using a hybrid of randomized response technique and crosswise design. Archives of sexual behavior, 46(2), 373-384. https://doi.org/10.1007/s10508-016-0740-4
- van Vliet, N., Nebesse, C., & Nasi, R. (2015). Bushmeat consumption among rural and urban children from Province Orientale, Democratic Republic of Congo. Oryx, 49(1), 165-174. https://doi.org/10.1017/S0030605313000549
- Warner, S. L. (1965). Randomized response: a survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 60, 63-9. https://doi.org/10.1080/01621459.1965.10480775
- Wolter, F. (2019). A New Version of the Item Count Technique for Asking Sensitive Questions: Testing the Performance of the Person Count Technique. Methods, data, analyses: a journal for quantitative methods and survey methodology (mda), 13(1), 169-192. https://doi.org/10.12758/mda.2018.04
- Wu, Q., & Tang, M. L. (2016). Non-randomized response model for sensitive survey with noncompliance. Statistical methods in medical research, 25(6), 2827-2839. https://doi.org/10.1177%2F0962280214533022
- Yu, J. W., Tian, G. L., & Tang, M. L. (2008). Two new models for survey sampling with sensitive characteristic: design and analysis. Metrika, 67(3), 251. https://doi.org/10.1007/s00184-007-0131-x