Experimental Validation of Specialised Questioning Techniques in Conservation

Ibbett, Harriet; Dorward, Leejiah; Dwiyahreni, A. A.; Jones, J.P.G.; Kaduma, J.; Kohi, E. M.; Mchomvu, J.; Prayitno, K.; Sabiladiyni, H.; Sankeni, S.; Saputra, A. W.; Supriatna, J.; St John, Freya A. V.

Conservation Biology

DOI:
10.1111/cobi.13908

Published: 01/10/2022

Peer reviewed version

Experimental Validation of Specialised Questioning Techniques in Conservation

Abstract

Conservation increasingly relies on social science tools to understand human behaviour. Specialised Questioning Techniques (SQTs) are a suite of methods designed to reduce bias in social surveys, and are widely used to collect data on sensitive topics, including compliance with conservation rules. Most SQTs have been developed in western, industrialised, educated, rich and democratic countries (so called WEIRD contexts), meaning their suitability in other contexts may be limited. Whether these techniques perform better than conventional direct questioning is important for those considering their use. Here, we adopt an experimental design to validate the performance of four SQTs (Unmatched Count Technique, Randomised Response Technique, Crosswise model, Bean method) against direct questions when asking about a commonly researched sensitive behaviour in conservation, wildlife hunting. We developed fictional characters, and for each method, asked respondents to report the answers that each fictional character should give when asked if they hunt wildlife. With data collected from 609 individuals living close to protected areas in two different cultural and socio-economic contexts (Indonesia, Tanzania), we quantified the extent to which respondents understood and followed SQT instructions and explored the socio-demographic factors that influenced whether they provided a correct response. Participants were more likely to refuse SQTs than direct questions and modelling suggested SQTs were harder for participants to understand. Demographic factors, including age and education level significantly influenced response accuracy. When sensitive responses were required, all SQTs (excluding Bean method) outperformed direct questions, demonstrating that SQTs can successfully reduce sensitivity bias. However, when asked about each method, most respondents (59-89%) reported they would feel uncomfortable using them to provide information on their own hunting behaviour, highlighting the considerable challenge of encouraging truthful reporting on sensitive topics. This work demonstrates the importance of assessing the suitability of social science methods prior to their implementation in conservation contexts.

Introduction

Theories, frameworks and tools from the social sciences are increasingly integrated into conservation research and practice (Bennett et al. 2016). With this transition comes a responsibility to critically examine the tools adopted to ensure they are fit for purpose. Many of the social science methods used in conservation have been developed in Western, Educated populations in Industrialised, Rich and Democratic contexts (so-called WEIRD populations; Henrich et al. 2010). However, cultural, sociological and psychological differences mean that methods and understandings developed in one context may be inappropriate when applied in another, with subsequent implications for data reliability and validity (Henrich et al. 2010). Assessing the relevance of methods when delivered in contexts different from those in which they were developed is thus of critical importance to those considering their use.

Questionnaires asking respondents directly about their beliefs, attitudes and behaviour, are commonly used to collect data within conservation, but data can suffer from bias, particularly if the research topic is sensitive (Nuno & St John, 2015). Respondents may fear repercussions if they reveal the truth, and thus censor their responses (sensitivity bias), or refuse to answer whole or parts of surveys altogether (non-response bias) (Blair et al. 2020). Developed by social scientists to overcome these biases, Specialised Questioning Techniques (SQTs) are increasingly applied in conservation to investigate illegal behaviours (Hinsley et al. 2018; Ibbett et al. 2021). Through varied mechanisms, SQTs ensure incriminating answers cannot be linked to individuals. Prevalence is estimated at the population level, and multivariate analyses can be applied post-hoc to identify characteristics of those possessing sensitive attributes (St John et al. 2012; Nuno & St John 2015). Compared to conventional questioning techniques (hereafter direct questions), SQTs are hypothesised to provide respondents greater protection, encourage more honest responding, and increase data accuracy (Chaudhuri & Christofides 2013). However, SQTs require careful design (Hinsley et al. 2018; Ibbett et al. 2021), are more complex to administer, and are less efficient as noise introduced by anonymising processes mean more data (and thus more resources) are needed to achieve SQT estimates with similar confidence to direct questions (Lensvelt-Mulders et al. 2005a).

Numerous SQTs exist, each developed to overcome the limitations of others (Nuno & St John 2015; Cerri et al. 2021). Some rely on probability to determine how respondents should answer. For example, Randomised Response Techniques (RRTs) use randomisers (e.g. dice) to determine whether a respondent should answer truthfully or provide a prescribed response (Ibbett et al. 2021). Other methods mask responses by aggregating answers. For example, the Unmatched Count Technique (UCT) divides the sample into two, one half are provided a list of innocuous items, the other receives the same list with the sensitive attribute added (Droitcour et al. 1991); respondents report how many of the listed items apply to them. The Crosswise model presents participants with one innocuous question with known prevalence and one question which is sensitive. Respondents report if their answer is the same for both questions, or yes to one question (Yu et al. 2008; Sagoe et al. 2021). Finally, developed for lower education contexts, and with reduced complexity compared to other SQTs (Lau et al. 2011), the Bean method asks respondents to secretly move specific-coloured beans from one jar to another, depending on their answer (Jones et al. 2020). See Appendix 1 for applications in conservation.

Whether SQTs reduce biases relative to direct questions is of critical importance to those designing surveys investigating sensitive topics. Ideally, the performance of SQTs is assessed by validating estimates against data on the true prevalence of the sensitive characteristic. However, difficulties associated with obtaining data on true prevalence means validation studies are rare (Blair et al. 2015). Indeed, a review of 35 years' of RRT research identified only six studies across multiple disciplines (Lensvelt-Mulders et al. 2005b). In the only validation study in conservation, Bova et al. (2018) covertly observed recreational anglers in South Africa and invited those who had been recorded breaking regulations to participate in a survey on angling compliance. Although all were observed breaking rules, only 79.6% of respondents admitted violations when asked to self-complete a questionnaire and deposit it in a ballot-box. Estimates from those surveyed face-to-face using direct questions or RRT were substantially lower (46.5% and 38.5% respectively). Other studies document similar findings (Wolter & Preisend¨orfer 2013; Rosenfeld et al. 2016), highlighting that while SQTs can reduce bias, their performance varies, and may still under estimate prevalence.

In lieu of being able to validate estimates against true prevalence, researchers commonly compare estimates derived from SQTs against direct questions, with the method that produces the highest estimate considered the most accurate and least biased (Blair et al. 2015). Numerous studies across disciplines demonstrate that SQTs perform better than direct questions when investigating a range of sensitive topics (e.g. Anglewicz et al. 2013; Stubbe et al. 2014), however, a substantial proportion also report the opposite (Coutts & Jann 2011; Höglinger et al. 2016), including in conservation science (e.g. Nuno et al. 2018; Davis et al. 2019). Whilst such findings can occur if the behaviour is exceptionally rare (St John et al. 2018; Ibbett et al. 2019), SQTs also have higher cognitive load (Solomon et al. 2007), are harder to understand (Coutts & Jann 2011; Davis et al. 2019), take longer to complete (Bova et al. 2018), and can arouse suspicion among respondents (Razafimanahaka et al. 2012). To be successful, SQTs require respondents to understand what they must do and why, and be willing to follow procedures fully (Hoffmann et al. 2017).

Recently, several experimental studies have contributed evidence on what affects how well SQTs work. For example, to explore how randomisers, phrasing of instructions and response options affect respondents' willingness to follow RRT instructions, John et al. (2018) conducted a series of online experiments. Similarly, to experimentally measure respondents' comprehension of five SQTs, Hoffmann et al. (2017) presented participants with descriptions of fictional characters, some who possessed the sensitive attribute (exams cheating), some who did not. Using each method, respondents reported the answer each fictional character should give when asked if they cheated in exams. How well respondents understood the method was calculated per respondent as the percentage of correct answers provided across all fictional characters. All SQTs were less comprehensible than direct questions, with less-educated respondents experiencing greater comprehension difficulties. While these studies provide invaluable insights into the efficacy of SQTs for asking sensitive questions, they were conducted in so called 'WEIRD' contexts (Henrich et al. 2010) and mostly online. Yet, due to various factors (e.g., lower literacy; poor technological access), conservation social science studies are often delivered face-to-face. Thus, understanding how SQTs perform under such conditions is crucial.

Here, we build on the experimental design of Hoffmann et al. (2017), adapting it to explore the performance of SQTs when asking people living around protected areas about a commonly researched sensitive behaviour, wildlife hunting. We collect data in-person in Indonesia and Tanzania; two non-WEIRD countries which are highly biodiverse, but significantly different in cultural and socio-economic terms. We aim to quantify the extent to which respondents understand and follow SQT instructions, and explore how socio-economic characteristics (age, gender and education) affect whether individuals answer correctly. Alongside conventional direct questioning, we test four SQTs, two frequently applied in conservation research, UCT and RRT (Hinsley et al. 2018; Ibbett et al. 2021), and two considered simpler to understand, but which are not yet widely applied in conservation, the Bean method (Jones et al. 2020) and Crosswise model (Yu et al. 2008).

Methods

Study sites

Data were collected from a selection of villages situated around the Leuser Ecosystem in northern Sumatra, Indonesia and the Ruaha-Rungwa protected area complex in Tanzania (Fig. 1). Both landscapes are sites of global conservation importance (Dickman et al. 2014; Myers et al. 2020), and include protected areas which restrict and regulate natural resource use. Hunting of any protected species, or hunting without a permit (unless for traditional use) is prohibited in Indonesia, while hunting any wild animal without permission is forbidden in Tanzania. Illegal hunting has been identified as a conservation concern at both sites (Pusparini et al. 2018; Beale et al. 2018), but has been little researched (although see Knapp et al. 2017; Hariohay et al. 2019) with no known applications of SQTs in either landscape. Previous studies have used UCT to investigate hunting elsewhere in Tanzania (Nuno et al. 2013; Wilfred et al. 2019), while RRT has been deployed, but with limited success in Indonesia (St John et al. 2018).

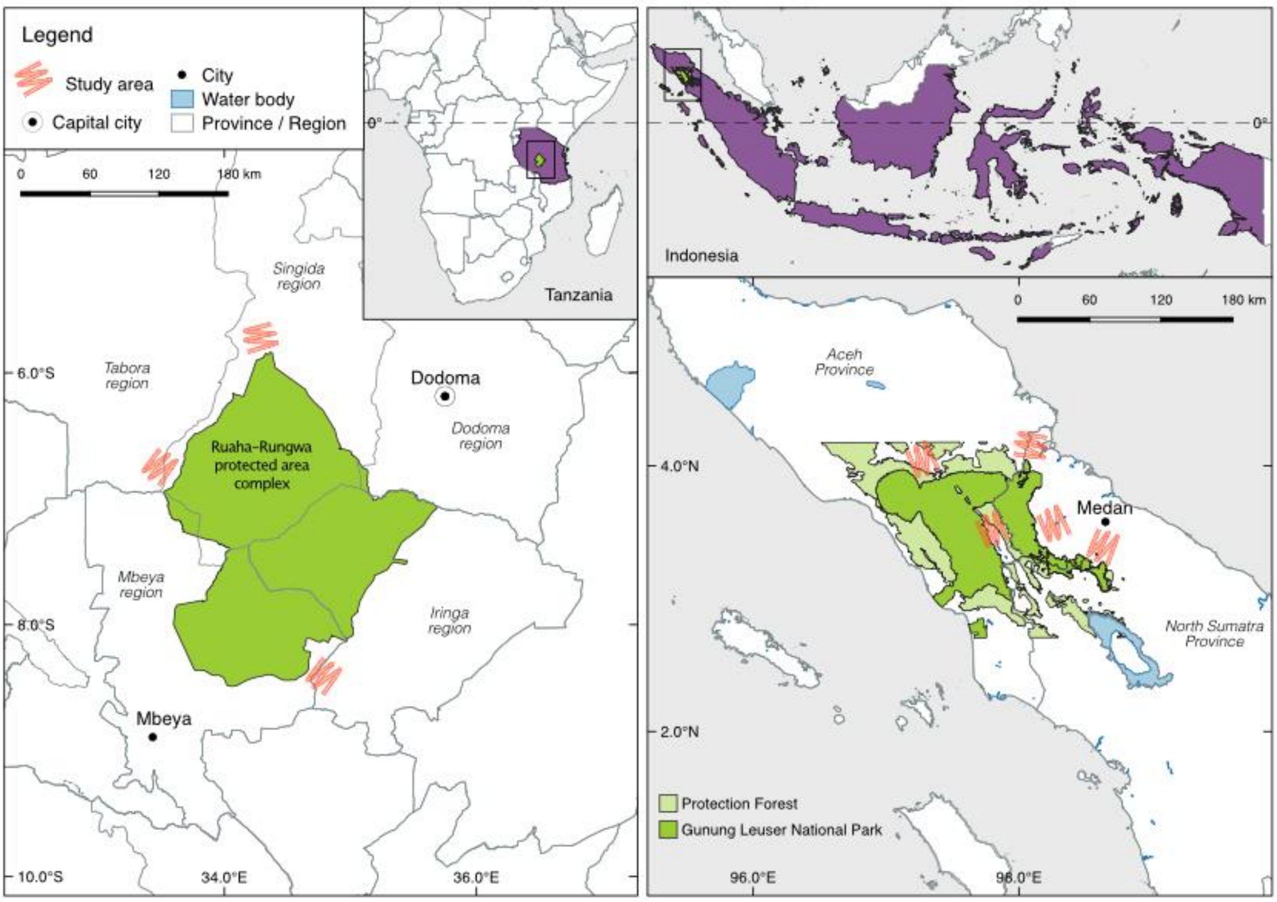

Figure 1. Data were collected in villages situated in two protected landscapes: the Leuser Ecosystem in northern Sumatra, Indonesia (seven villages), and the Ruaha-Rungwa protected area complex in southern Tanzania (six villages). In accordance with ethics approval, we do not indicate the precise locations of study villages.

Experimental design

Building on the experimental design of Hoffmann et al. (2017), we presented respondents with cards depicting fictional characters. Respondents were asked to imagine they were each of the fictional characters, and in turn for each method, answer questions about whether each fictional character hunted wildlife. Because the behaviour of each character was known, we could validate whether a respondent provided the correct answer and use this as a proxy to measure whether they understood, and followed, the instructions associated with each method.

Fictional characters

Five fictional characters were introduced to respondents using character cards (Appendix 1). The cards detailed information on the characters' birth month alongside four livelihood activities the character conducted (Fig. 2). Three characters conducted a sensitive activity (hunting wildlife), two did not. Characters served different purposes; Character One was used to introduce the method to respondents and Character Two was used to practice the method. We only proceeded to characters Three, Four and Five, once we were certain respondents understood instructions associated with each method. Characters Three to Five collected data on whether respondents provided accurate answers for each questioning method. To minimise respondent fatigue whilst maximising data on how respondents answered sensitive questions, two of these characters hunted, one did not. The order each character was presented to respondents was randomised, to eliminate order effects.

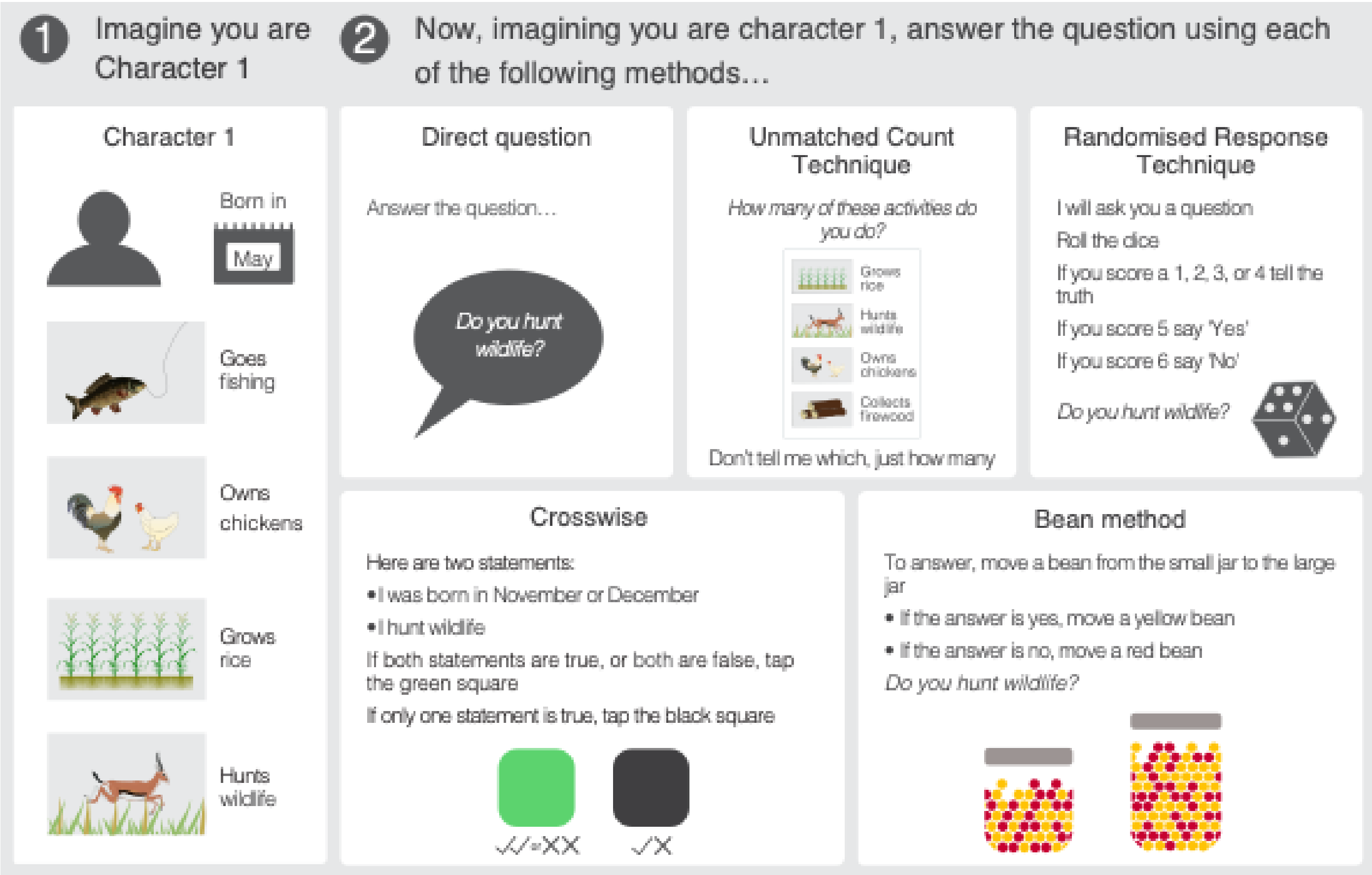

Figure 2. Example of a fictional character card (Step 1, left), and the instructions associated with each of the methods tested (Step 2, right). For each method, respondents were given detailed instructions on how to answer, and then asked to identify the answer the character should provide. Note the diagram only shows the RRT-dice method, for a description of the RRT-button method see Appendix 1.

Methods tested

Respondents received instructions for each method. In the direct question treatment, respondents were simply asked to provide a direct response of 'yes' or 'no', to the question about whether the character hunted wildlife. For the RRT, each respondent was provided a six-sided die in an opaque cup and asked to shake it without revealing the result to the interviewer. In Tanzania, if a 1 was rolled respondents were 'forced' to answer 'yes' regardless of whether this was true for the character. If a 2 was rolled respondents were 'forced' to answer 'no', and if a 3, 4, 5 or 6 was rolled, respondents were instructed to answer truthfully about the character's behaviour (Appendix 1). In Indonesia, the response options were reversed (i.e., 1, 2, 3, 4 =truthful, 5=forced-yes, 6=forced-no) to assess whether the order of forced responses affected performance. A die was chosen as the randomiser as they are commonly used in conservation RRT studies (Ibbett et al. 2021), and previous research conducted in similar contexts suggests they are effective (St John et al. 2015; Ruppert et al. 2020). Further, because evidence suggests randomiser choice can impact respondent's willingness to engage with the method (Coutts & Jann 2011; Razafimanahaka et al. 2012), in

Indonesia, we tested another randomiser; a cloth bag containing 8 orange buttons, 2 yellow buttons and 2 white buttons. Respondents were instructed to provide a truthful answer if an orange button was selected, to answer 'yes' if a yellow button was selected, and 'no' if a white button was selected (Appendix 1).

To test UCT, respondents were shown a card depicting four activities including hunting wildlife and asked to report the number of activities that applied to the fictional character (Appendix 1). When designing a UCT, researchers must be careful to avoid design effects which can occur if respondent report that all (ceiling effect), or none (floor effect), of the items apply to them (Droitcour et al. 1991), meaning careful piloting of UCT items is required (Hinsley et al. 2018). Our UCT design ensured that respondents were never required to report that a character conducted zero or four activities, thus avoiding ceiling and floor effects.

To respond using the Bean method, respondents were presented with two jars; one large, one small, and asked to secretly move a maize kernel from the small to large jar if the fictional character hunted wildlife, or a kidney bean if the fictional character did not. Jars were shaken before and after use, and were opaque, so as not to reveal the colour of the bean moved. Due to the emergence of the Covid-19 pandemic and subsequent impracticalities associated with adapting the method for safe enumeration (e.g., sanitising beans between respondents was impractical, and using multiple sets of jars would have undermined the privacy of the method) we were unable to test the Bean method in Indonesia.

For the Crosswise model, we presented respondents with two questions 1) were you born in November or December? and 2) do you hunt wildlife? and asked them to report whether the characters answer would be the same for both questions, or yes to only one (Appendix 1). To date, most applications of Crosswise model have been online, meaning participants are able to read the question-and-answer options (Meisters et al. 2020). However, our survey was enumerated face-to-face, with question-and-answer options read out to respondents. Preliminary piloting suggested this was problematic, as respondents had to remember the instructions and both questions. To overcome this, whilst piloting we developed a prompt card which featured a green square featuring two ticks and two crosses underneath, and a black square featuring one tick and one cross underneath (Appendix 1). Respondents were asked to tap the green square if their response was the same for both questions, and the black square if their response was yes to only one question (Fig 1).

Data collection

Survey instruments were developed in English and translated into the national languages of Bahasa Indonesia or Kiswahili by two team members fluent in the respective language. An independent back-translation was used to check the initial translation's accuracy. Questionnaire refinement occurred over five-weeks, and coincided with training and piloting. Questionnaires were administered face-to-face by KP, HS and AWS in Indonesia and SS, JM, JK in Tanzania. All data were collected using Open Data Kit (Brunette et al. 2013) on encrypted mobile phones. We adopted a convenience sampling strategy, with respondents recruited with the assistance of local guides, based on availability. Wherever possible, the team targeted male respondents aged 18 to 55, as this is the demographic most likely to be involved in hunting (Hariohay et al. 2019), thus information on how well this group of respondents understood SQTs was of interest for future research on rule-breaking.

We captured basic demographic data (respondent age, gender, years of education), alongside data on birth month. Birth month is often used as an alternative statement in Crosswise model designs (Sagoe et al. 2021), or as a randomiser for RRT (Ibbett et al. 2021). Yet, in many contexts people do not know their Gregorian birth date, therefore, it was important to determine how prevalent knowledge of birth month was, so that we could assess its feasibility as an alternative statement in future research.

We recorded the number of times participants practised each method using Character Two, before asking three questions using the method (using Characters Three, Four and Five; Appendix 1). For responses to direct questions, UCT and Crosswise model it was possible to immediately assess whether the respondent provided the correct response (because the answer was fixed). However, for RRTs and the Bean method, where responses depended on the outcome of a randomising event (i.e., dice roll or button selection) or movement of a bean, we were unable to verify if the respondent had provided the correct answer. Thus, after each RRT and bean question, respondents were asked to report the outcome of the randomising device (e.g., the number rolled, the button colour selected), or the type of bean moved. After each question, respondents were asked to rate, on a 5-point Likert scale, how much privacy they felt the method afforded. Five-point Likert-scales were also used to measure how well respondents felt they understood the method; how easy the method was to comprehend; how much protection respondents felt the method offered; and how comfortable respondents would feel providing honest responses about their own hunting behaviour using the method. For full methods see Appendix 1.

Ethical considerations

All data were anonymous. The study did not collect any sensitive data, as respondents were only asked about the rule-breaking behaviour of fictional characters. All respondents were over 18 years old and verbal consent was sought before every interview. As a token of thanks, participants were given a small, culturally appropriate gift. Research was formally approved by Bangor SNS Ethics Committee (coses2019hi01). HI and LJD accompanied SS, JM and JF in Tanzania throughout data collection (September-December 2019), but were unable to do so in Indonesia due to pandemic-related travel restrictions (data collected August-November 2020). Rigorous health and safety measures were implemented to mitigate Covid-19 transmission in survey communities. Research was conducted with the permission of national and local authorities.

Analysis

We performed analyses in R version 3.6.2 (R Core Team 2019). For each method, we calculated the percentage of correct responses per respondent, across all fictional characters. We then used descriptive statistics to explore data, assess respondent's understanding of methods and compliance with instructions, and test for collinearity between predictors prior to modelling. To examine which factors influenced whether a respondent answered a question correctly, we fitted generalised linear mixed models to each country dataset using the package lme4 (Bates et al. 2015). The response variable was a binary indicator of whether a respondent gave the correct answer to each question (Table 1). Respondent gender, age, years of education, the method tested, the number of practices required, interviewer and whether a sensitive response was required (i.e., the character hunted) were all included as fixed effects. We included interactions between method and whether a sensitive response was required, and method and years of education. Following Gelman & Hill (2007), to improve the interpretability of coefficients, continuous variables for respondent age, years of education and the number of practices were scaled and centred by subtracting the mean and dividing by two standard deviations. Random effects were included to control for respondent and method. To achieve convergence, models were fitted using a BOBYAQA optimizer, and tested for singularity. Models showed no significant signs of dispersion when checked using the DHARMa package (Harting 2020). Tukey post-hoc tests were conducted to assess pairwise correlations between each method.

Table 1. Explanation of the response and predictor variables tested in country-specific binomial general linear mixed models

Variables	**Description *(Data type: Levels)***
Response variable	
Did the respondent provide the correct answer?	*(Categorical: Yes / No)*
Predictor variables *(Effect type)*	
ID *(Random effect)* [a]	Unique ID code assigned to each respondent *(Continuous)*
Age	Age of respondents in years *(Continuous)*
Gender	Gender of the respondent *(Categorical: Male / Female)*

Education	Number of years schooling the respondent completed *(Continuous)*
Method *(Random effect)* [b]	The method tested (*Categorical: Direct question / UCT / RRT-dice / RRT-button / Crosswise model / Bean method)*
Practices	The number of practices the respondent required before providing the correct response *(Continuous)*
Interviewer	ID of the interviewer administering the questionnaire *(Categorical: One, Two, Three)*
Response sensitive	Whether a sensitive response was required (i.e., whether the respondent was required to report that a character hunted) *(Categorical: Sensitive / Non-sensitive)*
Interaction terms:	
Method * Response sensitive	
Method * Education	

[a] Included as a random effect to control for respondents answering multiple questions per method
[b] Included as a random effect to control for one question being asked for each of the three characters, per method

Results

Respondent demographics

Data were collected from 303 people in Indonesia and 306 in Tanzania. The gender of both samples was biased towards men (Indonesia, 75% male, Tanzania, 56%). Education levels were higher in Indonesia (mean 9.9 years, ±SE0.207), than Tanzania (6.6 years, ±SE0.180). In Tanzania men had significantly more years of education than women (7 and 6 years respectively, t=-2.864, df=280, p=0.005). There was no relationship between gender and education in Indonesia (9.9 years, t=0.278, df=116, p=0.781). The mean age of respondents sampled in both countries was 38 years (Indonesia, min:18, max:60, SE0.752; Tanzania, min:18, max:80, SE0.569). Most respondents knew their birth month (Indonesia, 83.5%; Tanzania, 73.5%).

Non-response

Levels of non-response varied by method and country. Overall, questions were refused more often in Tanzania than Indonesia. In both countries, RRT-dice was the method most frequently refused (Table 2), followed by Crosswise model. Direct questions received the least refusals in both countries, followed by UCT.

Table 2. Sample sizes and numbers of non-response per method in each country. Each method was repeated three times, per respondent.

Method	Indonesia (N=303)		Tanzania (N=306)	
	Responses	Refusals	Responses	Refusals
Direct questions	909	0	842	76 (8%)
UCT	908	1 (<1%)	798	120 (13%)
Crosswise model	904	5 (1%)	767	151 (16%)
RRT-dice	891	18 (2%)	761	157 (17%)
RRT-button	909	0	-	-
Bean method	-	-	784	134 (14%)
Total	4521	27 (<1%)	3,952	638 (14%)

Responses = number of questions answered per method; Refusals = number of questions refused per method.

Percentage of correct responses per method (by country)

In Indonesia, UCT and direct questions resulted in the highest percentage of correct responses (90.1%, 95%-CI±1.9% and 89.4%±2.0% respectively; Fig. 3); fewer correct responses were reported via RRT-dice and RRT-button (dice:

81.0%±2.6%, button: 82.8%±2.4%), while Crosswise model resulted in the lowest percentage of correct responses (64.3%±3.1%). In Tanzania, RRT-dice, UCT and direct questions secured the highest percentage of correct responses (80.0%±2.8%, 78.9%±2.8% and 77.2%±2.8% respectively; Fig. 3), in comparison the Bean method and Crosswise model both performed significantly worse (67.6%±3.3% and 65.0%±3.4%).

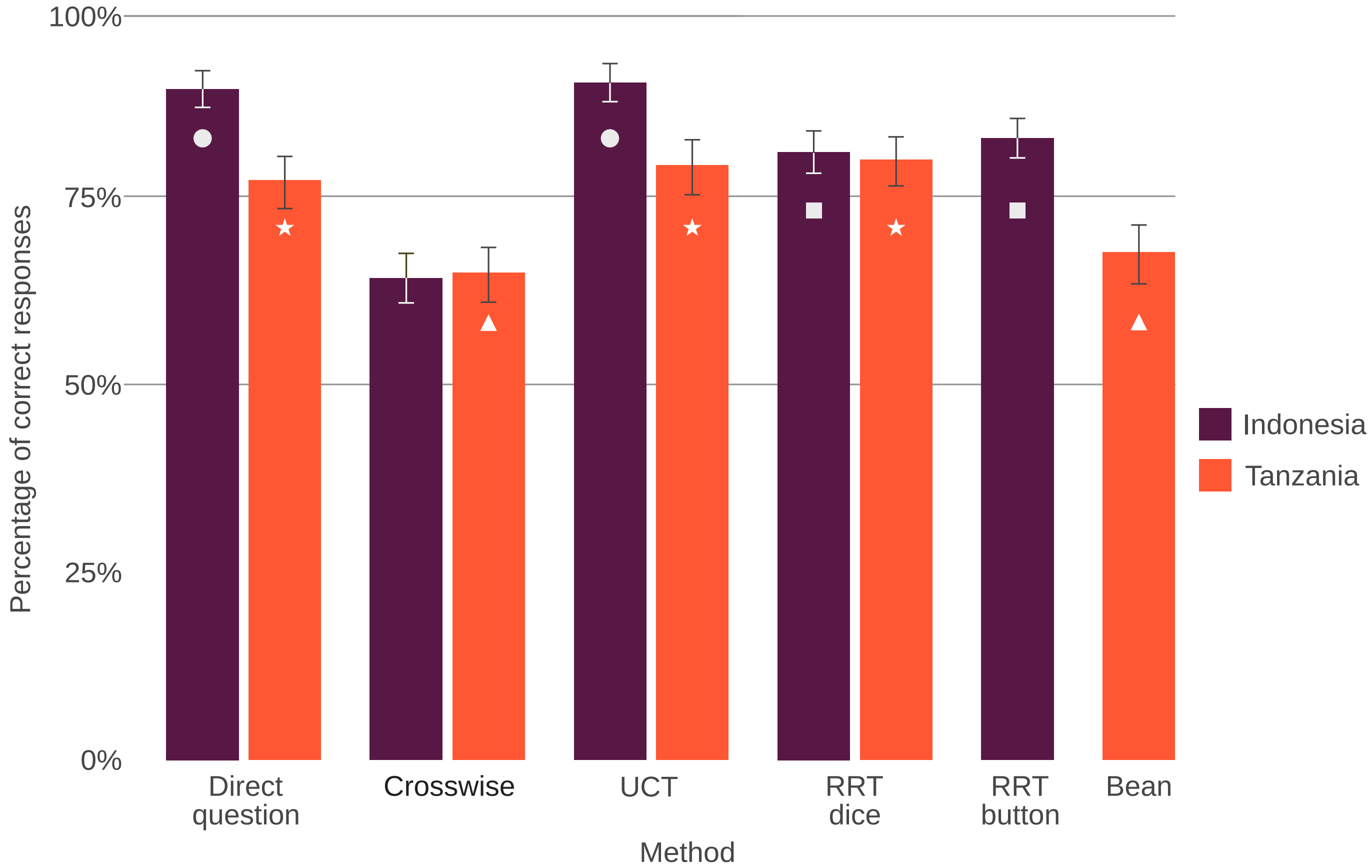

Figure 3. Mean percentage of correct responses for each method tested in Indonesia (purple bars) and Tanzania (orange bars). Error bars represent 95% confidence intervals. Matching shapes indicate there was no significant difference in the mean percentage of correct responses between these methods when tested in the same country.

Socio-demographic predictors of correct responses

Modelling showed several factors predicted whether a respondent answered correctly (Figs. 4, 5; Appendix 3). In Indonesia, women were more likely than men to answer correctly, although there was no effect of gender in Tanzania. In both countries, the likelihood of a correct response decreased with age. Education was not a significant predictor of a correct response in Indonesia, but in Tanzania the more years of education a respondent had completed, the greater the probability they would answer correctly. Results suggest respondents who required more practices were also more likely to answer incorrectly, with more practices required on average per respondent in Tanzania, than Indonesia. Who was delivering the survey impacted response accuracy in Indonesia, with respondents questioned by interviewer two, significantly less likely to answer correctly.

The impact of method & response sensitivity on whether respondents answered correctly

According to our model, respondents were more likely to provide a correct response when answering a direct question, compared to all other methods (Fig. 4, 5, Appendix 3), although in Tanzania, direct questioning did not perform significantly better than RRT-dice or UCT (Tukey post-hoc tests, Appendix 3). When compared against each other, all SQTs performed equally, with the exception of the Crosswise model, which performed significantly worse than other SQTs (Appendix 3). Whether the character hunted, and thus whether the respondent was required to provide a sensitive response, was a significant predictor of whether a respondent answered correctly. In both countries participants were less likely to provide correct answers when the character hunted. Findings suggested a significant interaction between method and whether a sensitive response was required. When respondents were required to provide a sensitive answer (i.e., where the character card depicted that the character hunted), the probability of a respondent providing a correct response was significantly higher if answering using an SQT compared to direct questions. This applied in both countries and for all SQTs except the Bean method, suggesting that, except for the Bean method, SQTs outperformed direct questions when a socially undesirable response was required (Appendix 3). This effect was particularly pronounced for Crosswise model which demonstrated the greatest difference in probability of a correct answer when a sensitive response was and was not required. Results indicate that there was little overall interaction between education and method, except in Tanzania where those who had more years of education were less likely to provide a correct response using Crosswise model (Appendix 3).

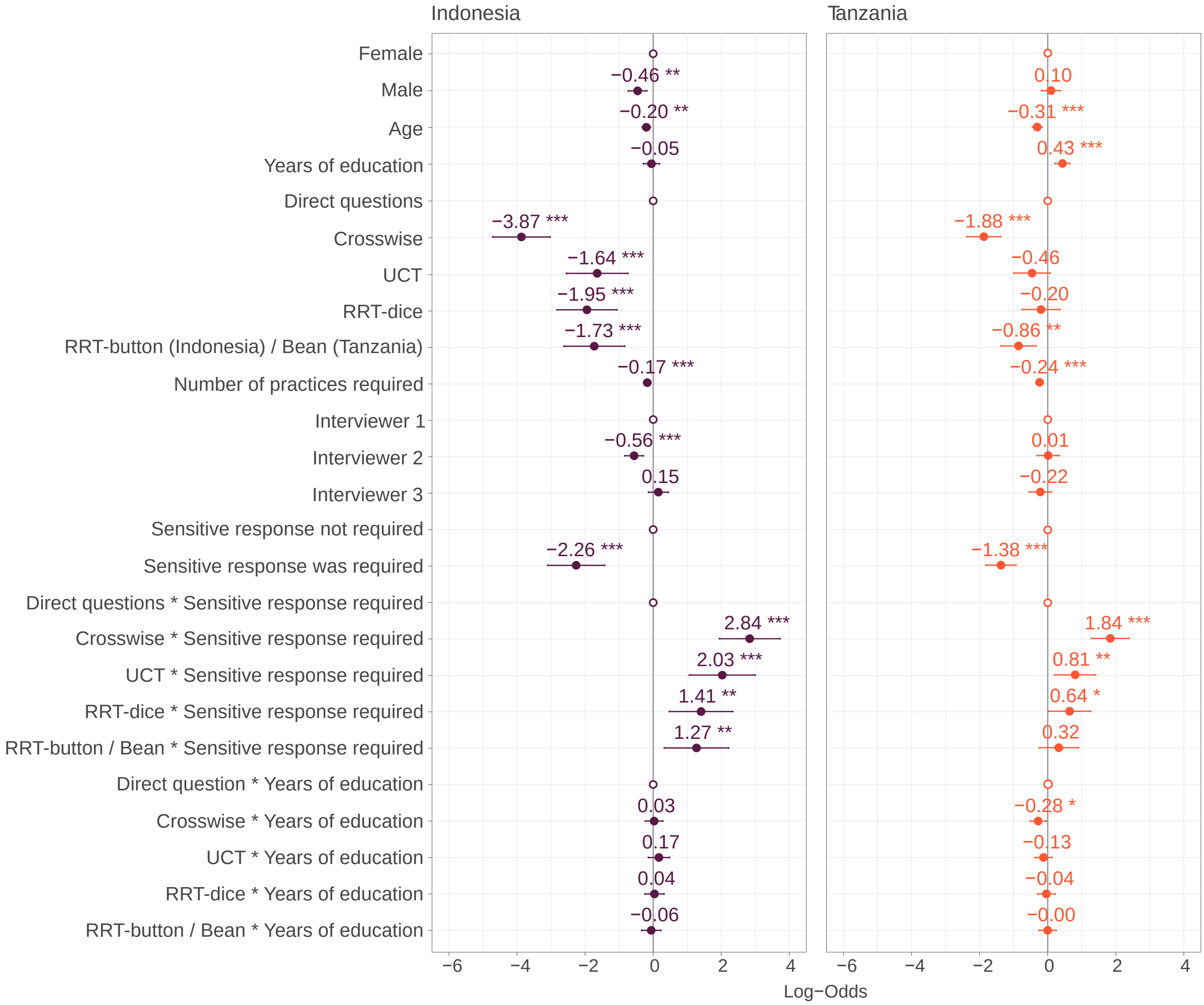

Figure 4. Log-odds regression coefficients with standard errors from a binomial general linear mixed model, with random effects for respondent and method, where the binomial response variable represents whether the respondent answered the question correctly or not. Note, the RRT-button was applied in Indonesia only, and the Bean method was applied in Tanzania only. Circles with white centres represent reference categories for categorical variables. Significance codes: '***'<0.001, '**'<0.01, '*'<0.05

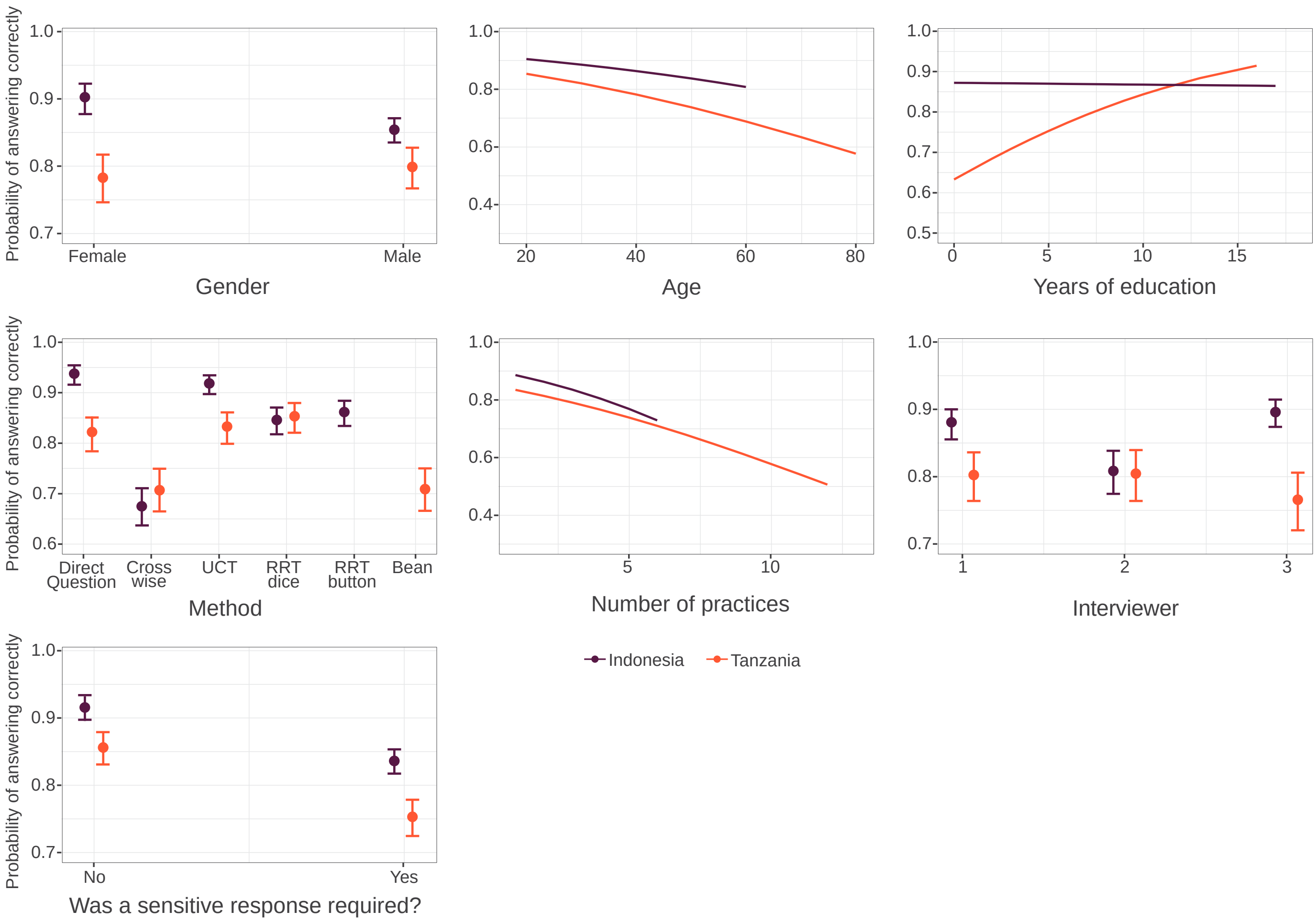

Figure 5. Marginal-effects plots for each fixed-effect included in the model, showing the probability of answering correctly. Error bars represent 95% CIs.

Compliance with instructions

For both RRTs, whether the respondent gave the correct response was verified by the respondent providing information on their action (e.g., die number rolled or button selected) (Appendix 3). Analysis of these data is useful for assessing whether respondents complied with the instructions provided. For a standard six-sided die, the probability of rolling any number is 0.167, meaning all numbers (one to six) should have been reported in equal abundance. For the RRT-button, the probability of selecting an orange button was 0.66 and the probability of selecting a white or a yellow button was 0.17. In Tanzania, the number of times each dice number was reported was significantly different to expected (χ^2=28.658, df=5, p-value=<0.001); respondents over-reported options which instructed them to give forced responses and under reported responses that required truthful answers. While a similar trend was seen in Indonesia, it was not significant for either the RRT-dice (χ^2=7.162, df=5, p-value=0.209), or RRT-button (χ^2=5.806, df=2, p-value=0.055; Appendix 3).

Respondent's self-reported understanding of methods

Direct questions were generally considered easier to understand than SQTs (Table 3). In Indonesia, 90% of respondents found direct questions easy or very easy to answer, while in Tanzania, UCT was considered easiest to answer (82% of respondents, although this was only marginally more than direct questions (79%); Table 3). Overall, few respondents reported that they would feel comfortable providing honest responses about their own hunting behaviour using any of the methods, especially in Tanzania. However, people reported that they would be more comfortable using SQTs than direct questions. Similarly, in both countries, a higher percentage of respondents felt that SQTs kept their answers secret or very secret compared to direct questions. Respondents reported understanding direct questions better than any of the SQTs (except UCT in Tanzania; Table 3). Crosswise model was reported as least well understood in both countries.

Table 3. Percentage of respondents in Indonesia (N=303) and Tanzania (N=306) who reported agreement with each of these statements.

	Respondents who felt questions were easy or very easy to answer using this method		Respondents who would feel comfortable or very comfortable providing honest responses about their own hunting behaviour using this method		Respondents who felt the method kept their answer secret or very secret		Respondents who understood or understood well the method	
	Indonesia	Tanzania	Indonesia	Tanzania	Indonesia	Tanzania	Indonesia	Tanzania
Direct questions	90%	79%	25%	11%	33%	44%	88%	92%
Crosswise model	59%	61%	32%	13%	41%	58%	56%	76%
UCT	86%	82%	36%	14%	50%	61%	78%	93%
RRT-dice	65%	73%	41%	15%	51%	64%	63%	85%
RRT-button	66%	-	31%	-	51%	-	62%	-
Bean method	-	80%	-	18%	-	62%	-	89%

Discussion

To develop effective interventions, conservationists require reliable information about human behaviour, including the proportion of a population engaged in illegal or otherwise sensitive behaviours (St John et al., 2013). Designed to reduce bias, Specialised Questioning Techniques are increasingly applied in conservation, but with mixed success (Cerri et al. 2021), leading researchers to question exactly how well research participants understand and follow SQTs instructions (Hinsley et al. 2018; Davis et al. 2019). Conservation research is often conducted in different contexts and conditions to those in which SQTs were developed, meaning it is important to determine how factors such as education level, gender and face-to-face enumeration affect how well respondents understand SQTs, and how comfortable respondents feel using these methods. Here, we provide valuable insights for conservationists considering SQT use in the field.

In both Indonesia and Tanzania, the probability of a respondent answering an SQT correctly was lower than for direct questions, suggesting they were harder for respondents to understand (Hoffmann et al. 2017; Davis et al. 2019), particularly in Tanzania, if education level was low. This is likely because SQTs involve more instructions, and often rely on the use of additional equipment such as dice, beans and jars, or lists. Together, these factors increase cognitive load, making it harder for respondents to follow instructions (Hoffmann et al. 2020). With similar findings from two culturally distinct countries we can make general recommendations about the contexts in which SQTs might be better understood by respondents. For example, respondents were more likely to answer correctly about a fictional character's behaviour, and thus to have understood instructions, when they had more years of education, and they were younger. Another good indicator of respondent understanding was the number of practices required; the more an interviewer had to explain a method, the lower the likelihood instructions were understood. Therefore if, when piloting, excessive explanation is required to introduce the method to participants, researchers should consider whether the method is appropriate, and if many explanations are required, interpret data cautiously. Findings also reinforce that *who* asks questions matters. It is important to consider how factors including interviewer characteristics (e.g., gender, age, manner and personality) influence research, being mindful that interviewers vary in their experience, how comfortable they make respondents feel and the quality of data they collect (Blair et al. 2020).

Overall, respondents' understanding of SQTs varied across the methods tested. Estimates from the percentage of respondents answering correctly and respondents self-reported evaluation of each method suggests UCT was the SQT understood best in both countries. While this infers some superiority compared to other SQTs tested, pairwise comparisons show that UCT was not better understood relative to other SQTs (excluding Crosswise model). Complexities associated with the selection of list items mean UCTs may not always be an appropriate or feasible method, particularly if asking about multiple behaviours (Hinsley et al. 2018), or if low prevalence is expected (Ibbett et al. 2019; Davis et al. 2020). The finding that Crosswise model was poorly understood was surprising as other studies suggest it is easier to comprehend than alternatives (Hoffmann et al. 2017; Höglinger & Jann 2018). However, previous studies relied on self-administration, either online or using printed questionnaires (Sagoe et al. 2021), whereas our respondents relied on verbal instructions, meaning respondents had to remember instructions and questions. There is evidence that Crosswise model has a tendency to produce false-positive responses, leading to overestimations of prevalence (Höglinger et al. 2016; Höglinger & Jann 2018), although Meisters et al. (2020) suggest this can be overcome by providing respondents with more comprehensive and detailed instructions. While Crosswise model shows potential where self-administration is viable, the low overall comprehension that we detected suggests significant adaptation is required to deploy this method face-to-face, particularly in lower literacy contexts where written instructions may not be appropriate. More surprising is how poorly the Bean method performed. Promoted for its ease of use, particularly in low literacy contexts, the method involves clear, simple instructions and relies on familiar equipment; dried beans and jars (Jones et al. 2020). Yet, when tested in Tanzania we found the percentage of correct responses was relatively low, despite a high proportion of respondents reporting they found the method easy to use and that they understood instructions. Some of this error could be attributed to interviewers incorrectly counting beans (Jones et al. 2020) and the experimental nature of the exercise (having to report the behaviour of a character). Further error may also result from purposeful false responding; when asked how private they felt the method was, some respondents reported low levels of privacy, suggesting interviewers would look in the jar to determine what bean they had moved. One respondent suggested it was possible to satellite track the movement of individual beans, highlighting concerns about the trustworthiness of researchers, as well as the use of surveillance technologies in monitoring communities' activities (Sandbrook et al. 2021).

Ultimately, SQTs are designed to protect research participants when collecting sensitive data. While our study mostly provides insights into whether respondents understand and follow the instructions of different SQTs, we also investigated whether the likelihood of respondents answering correctly was influenced by whether they were required to provide sensitive responses (i.e., report that the fictional character hunted). When sensitive responses were required, all SQTs (except the Bean method) significantly increased the likelihood of respondents giving a correct response relative to direct questions. This is the result we would expect if respondents were answering about their own behaviour (and suggests that SQTs do reduce sensitivity bias). Interestingly, this result was observed even though respondents were answering on behalf of fictional characters. This effect was strongest for the Crosswise model, perhaps because in this method, there is no 'safe' response option; both answers can be chosen by those who do and do not possess the sensitive attribute (Hoffmann et al. 2020).

Although our findings suggest SQTs can reduce sensitivity bias, they may exacerbate other forms of bias, such as non-response and evasive response bias. All SQTs in both countries received higher refusals than direct questions, with RRT receiving the highest number of non-responses. When researching illegal consumption of bushmeat in Madagascar, Razafimanahaka et al. (2012) found respondents did not like being 'forced' to admit to eating certain species by the RRT design and therefore refused to answer. Moreover, responding can be affected by randomiser type. Although we found no effect of randomiser type in Indonesia, participants associated dice with gambling in both countries, while in Tanzania some participants refused to touch equipment, concerned that we were trying to con and/or curse them. Alternatively in some cultures, certain numbers are considered lucky or unlucky (e.g., Yang 2011), which might impact how people interact with number-dependent randomisers. Additionally, the order RRT response options are provided to respondents may influence answers, for example, respondents may fixate on the 'safest' or most desirable answer they hear (e.g., forced-no), and fail to listen to all options. While extensive piloting and adopting different RRT designs such as the unrelated-question design can overcome the likelihood of non-response bias (Ibbett et al. 2021), high refusals also emphasise wider issues regarding efficiency. Due to the additional noise introduced to the data by anonymisation processes, compared to direct questions, all SQTs require larger sample sizes, and thus more research resource (Hinsley et al. 2018; Ibbett et al. 2021).

To be successful, SQTs rely on the assumption that those who do not possess the sensitive trait will comply with instructions and respond appropriately (Krumpal & Voss 2020). However, Krumpal & Voss (2020) suggest methods such as RRT can actually enhance socially desirable responding rather than reduce it, particularly when those who do not possess the sensitive trait are forced to provide affirmative responses. Like Chuang et al. (2021), our data suggest that some respondents understood the instructions but deliberately chose not to comply with them, mostly when sensitive responses were required, and particularly for the bean and RRT methods. While a large number of false-negative responses may lead to underestimations of prevalence, false-positives (which can also occur if respondents deliberately choose not to follow instructions, or if they misunderstand them) can be just as harmful. For example, false-positives may lead conservationists to believe that prevalence is higher than reality, resulting in inappropriately targeted interventions. Several techniques have emerged to counter this. For example, internal consistency checks can be used to identify potential bias (Cerri et al. 2021; Chuang et al. 2021), while designs such as the double-list UCT (Glynn 2013), or cheating detection variant of the RRT (Clark & Desharnais 1998), can help quantify potential bias. However, so far, the potential of these approaches is largely theoretical, with few empirical examples of their effectiveness (Cerri et al. 2021).

Reliance on fictional characters to explore respondents understanding of methods had limitations. Firstly, the use of characters added complexity to the response process, which may have decreased overall understanding of the methods. Consequently, our estimates may only represent minimal levels of understanding per method. Conversely, as respondents were not required to provide information about their own behaviour, they may have been more willing to engage than they would be in a conventional survey. Our results showed that some respondents deliberately failed to comply with SQT instructions as they felt uncomfortable admitting to a fictional character conducting sensitive behaviours, suggesting that if applied to their own behaviour, there may have been more evasive responses or refusals. Skewed prevalence of hunting amongst characters may have also aroused suspicion and affected responding, as respondents were asked to report hunting more often than not. Moreover, our design involved considerable repetition, with the surveys ranging from 45 minutes to two hours in duration depending on the skill of the interviewer, the respondent and the interview environment. This became tedious for some respondents and may have resulted in bias, with individuals providing answers simply to finish sooner. Shortening the survey by adopting a block-experimental design could overcome this challenge, but potentially at the cost of participant intra-comparability. With any experiment, results should be considered cautiously and within the confines of their limitations.

Despite the significant ways that SQTs aim to minimise risk to respondents, our study highlights the substantial effect of sensitivity when conducting conservation research on illegal behaviours. Our respondents were never asked about their own behaviour, the experimental nature of the research was emphasised throughout, and respondents were only required to provide information on fictional characters, yet sensitivity still affected responses. Concern that answers would be used to incriminate individuals in hunting was particularly high in Tanzania, with some respondents associating the survey as a form of trickery, especially when combined with methods such as the RRT, which forced participants to provide undesirable responses. While research previously conducted in Ruaha-Rungwa successfully gathered qualitative information on hunting, data were only obtained after key informants encouraged other community members to approach researchers (Knapp et al. 2017). The concerns we encountered emphasise the complexity of relationships that exist between communities and conservation, especially around protected areas, where regulations restricting people's access to, and use of natural resources are often strongly enforced. Conservation research often occurs in contested spaces, and both the Ruaha-Rungwa protected area complex, and the Leuser Ecosystem have turbulent colonial histories associated with dispossession (Walsh 2007; Minarchek 2020). Researchers asking about wildlife or natural resource use in such places are rarely perceived as neutral parties and are often assumed to be affiliated with conservation NGOs, government or protected area management (Brittain et al. 2020). Distrust of researchers' intentions and their use of data is subsequently high. Not only does this raise ethical questions about whether methods, such as RRT, that "force" respondents to admit to illegal behaviours causing potential distress, are appropriate, but it emphasises the need for ethical procedures such as free, prior and informed consent, that promote transparency and awareness of the research objectives (Brittain et al. 2020). It also highlights the importance of embedding research within enduring conservation efforts (e.g., Ruppert et al. 2021), and practices such as disseminating research findings back to communities (Brittain et al. 2020).

While social science has made significant strides in developing methods that reduce bias during sensitive research, our study highlights that these methods are not understood by all respondents, and even if they are, respondents may not feel comfortable enough to provide honest responses. To be successful, conservation researchers must be sensitive to the context in which the research will occur, have awareness about how conservation is perceived by potential study participants, and should pilot extensively to develop a robust and appropriate study design. Fundamentally, our work demonstrates the importance of assessing the suitability of social science methods prior to their implementation in contexts that differ substantially from where they were developed, as cultural, sociological and psychological differences may have substantial implications on data reliability and validity.

Supporting Information

Additional information is available online in the Supporting Information section at the end of the online article. The authors are solely responsible for the content and functionality of these materials. Queries (other than the absence of material) should be directed to the corresponding author.

Literature Cited

Anglewicz P, Gourvenec D, Halldorsdottir I, O'Kane C, Koketso O, Gorgens M, Kasper T. 2013. The effect of interview method on self-reported sexual behavior and perceptions of community norms in Botswana. AIDS and Behavior **17**:674–687.

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software **67**:1–48.

Beale CM, Hauenstein S, Mduma S, Frederick H, Jones T, Bracebridge C, Maliti H, Kija H, Kohi EM. 2018. Spatial analysis of aerial survey data reveals correlates of elephant carcasses within a heavily poached ecosystem. Biological Conservation **218**:258–267.

Bennett NJ et al. 2016. Mainstreaming the social sciences in conservation. Conservation Biology **31**:56–66.

Blair G, Coppock A, Moor M. 2020. When to Worry about Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments. American Political Science Review **114**:1297–1315.

Blair G, Imai K, Zhou YY. 2015. Design and Analysis of the Randomized Response Technique. Journal of the American Statistical Association **110**:1304–1319.

Bova CS, Aswani S, Farthing MW, Potts WM. 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.

Brittain S et al. 2020. Ethical considerations when conservation research involves people. Conservation Biology **00**:1–9.

Brunette W, Sundt M, Dell N, Chaudhri R, Breit N, Borriello G. 2013. Open Data Kit 2.0: Expanding and refining information services for developing regions. Page ACM HotMobile 2013: The 14th Workshop on Mobile Computing Systems and Applications.

Cerri J, Davis EO, Veríssimo D, Glikman JA. 2021. Specialized questioning techniques and their use in conservation: A review of available tools, with a focus on methodological advances. Biological Conservation **257**:1–19.

Chaudhuri A, Christofides TC. 2013. Indirect Questioning in Sample Surveys, 1st edition. Berlin.

Chuang E, Dupas P, Huillery E, Seban J. 2021. Sex, lies, and measurement: Consistency tests for indirect response survey methods. Journal of Development Economics **148**:102582.

Clark SJ, Desharnais RA. 1998. Honest Answers to Embarrassing Questions: Detecting Cheating in the Randomized Response Model. Psychological Methods **3**:160–168.

Coutts E, Jann B. 2011. Sensitive questions in online surveys: Experimental results for the randomized response technique (RRT) and the unmatched count technique (UCT). Sociological Methods and Research **40**:169–193.

Davis EO, Crudge B, Lim T, O'Connor D, Roth V, Hunt M, Glikman JA. 2019. Understanding the prevalence of bear part consumption in Cambodia: A comparison of specialised questioning techniques. PLoS ONE **14**:1–17.

Davis EO, Gaffi L, Mussoni G, Zaw T, Glikman JA. 2020. Insights into medicinal wildlife consumption and bear part use in Rakhine, Myanmar. Journal for Nature Conservation **58**:125923.

Dickman AJ, Hazzah L, Carbone C, Durant SM. 2014. Carnivores, culture and "contagious conflict": Multiple factors influence perceived problems with carnivores in Tanzania's Ruaha landscape. Biological Conservation **178**:19–27.

Droitcour J, Caspar RA, Hubbard ML, Parsley TL, Visscher W, Ezzati TM. 1991. The item count technique as a method of indirect questioning: A review of its development and a case study application. Pages 185–210 in S. Biemer, P.P, Groves, R.M., Lyberg, L.E., Mathiowetz, N.A., Sudman, editor. Measurement Errors in Surveys. John Wiley & Sons, Inc., New York, NY.

Gelman A, Hill J. 2007. Data analysis using regression and multilevel/hierarchial models. Cambridge University Press, Cambridge, UK.

Glynn AN. 2013. What can we learn with statistical truth serum? Public Opinion Quarterly **77**:159–172.

Hariohay KM, Ranke PS, Fyumagwa RD, Kideghesho JR, Røskaft E. 2019. Drivers of conservation crimes in the Rungwa-Kizigo-Muhesi Game Reserves, Central Tanzania. Global Ecology and Conservation **17**:e00522.

Harting F. 2020. DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. Available from https://cran.r-project.org/package=DHARMa.

Henrich J, Heine SJ, Norenzayan A. 2010. The weirdest people in the world? Behavioral and Brain Sciences **466**:61–83.

Hinsley A, Keane AM, St John FAV, Ibbett H, Nuno A. 2018. Asking sensitive questions using the unmatched count technique: Applications and guidelines for conservation. Methods in Ecology and Evolution **00**:1–12.

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 **52**:1768–1782. Behavior Research Methods.

Hoffmann A, Waubert de Puiseau B, Schmidt AF, Musch J. 2017. On the comprehensibility and perceived privacy protection of indirect questioning techniques. Behavior Research Methods **49**:1470–1483. Behavior Research Methods.

Höglinger M, Jann B. 2018. More is not always better: An experimental individual-level validation of the randomized response technique and the crosswise model. PLoS ONE **13**:1–22.

Höglinger M, Jann B, Diekmann A. 2016. Sensitive questions in online surveys: An experimental evaluation of different implementations of the randomized response technique and the crosswise model. Survey Research Methods **10**:171–187.

Ibbett H, Jones JPG, St John FAV. 2021. Asking sensitive questions in conservation using Randomised Response Techniques. Biological Conservation **260**.

Ibbett H, Lay C, Phlai P, Song D, Hong C, Mahood SP, Milner-Gulland EJ. 2019. Conserving a globally threatened species in a semi-natural, agrarian landscape. Oryx **53**:181–191.

John LK, Loewenstein G, Acquisti A, Vosgerau J. 2018. When and why randomized response techniques (fail to) elicit the truth. Organizational Behavior and Human Decision Processes **148**:101–123.

Jones S, Papworth S, Keane AM, Vickery J, St John FAV. 2020. The bean method as a tool to measure sensitive behavior. Conservation Biology **0**:1–10.

Knapp EJ, Peace N, Bechtel L. 2017. Poachers and Poverty: Assessing Objective and Subjective Measures of Poverty among Illegal Hunters Outside Ruaha National Park, Tanzania. Conservation and Society **15**:24–32.

Krumpal I, Voss T. 2020. Sensitive Questions and Trust: Explaining Respondents' Behavior in Randomized Response Surveys. SAGE Open **10**.

Lau JTFF, Yeung NCYY, Mui LWHH, Tsui HY, Gu J. 2011. A simple new method to triangulate self-reported risk behavior data-the bean method. Sexually Transmitted Diseases **38**:788–792.

Lensvelt-Mulders GJLM, Hox JJ, van der Heijden PGM. 2005a. How to improve the efficiency of randomised response designs. Quality and Quantity **39**:253–265.

Lensvelt-Mulders GJLM, Hox JJ, van der Heijden PGM, Maas CJM. 2005b. Meta-analysis of Randomised Response Research: Thirty-five years of validation. Sociological Methods & Research **33**:319–348.

Meisters J, Hoffmann A, Musch J. 2020. Can detailed instructions and comprehension checks increase the validity of crosswise model estimates? PLoS ONE **15**:1–19.

Minarchek M. 2020. Creating environmental subjects: Conservation as counter-insurgency in Aceh, Indonesia, 1925–1940. Political Geography **81**:102189.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB da, Kent J. 2020. Biodiversity hotspots for conservation priorities. Nature **403**:853–858.

Nuno A et al. 2018. Understanding implications of consumer behavior for wildlife farming and sustainable wildlife trade. Conservation Biology **32**:390–400.

Nuno A, St John FAV. 2015. How to ask sensitive questions in conservation: A review of specialized questioning techniques. Biological Conservation **189**:5–15.

Nuno A, Bunnefeld N, Naiman LC, Milner-Gulland EJ 2013. A Novel Approach to Assessing the Prevalence and

Drivers of Illegal Bushmeat Hunting in the Serengeti. Conservation Biology **27**:1355–1365.

Pusparini W, Batubara T, Surahmat F, Ardiantiono, Sugiharti T, Muslich M, Amama F, Marthy W, Andayani N. 2018. A pathway to recovery: The Critically Endangered Sumatran tiger Panthera tigris sumatrae in an “in danger” UNESCO World Heritage Site. Oryx **52**:25–34.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, Vienna, Austria. Available from https://www.r-project.org/.

Razafimanahaka JH, Jenkins RKB, Andriafidison D, Randrianandrianina F, Rakotomboavonjy V, Keane A, Jones JPG. 2012. Novel approach for quantifying illegal bushmeat consumption reveals high consumption of protected species in Madagascar. Oryx **46**:584–592.

Rosenfeld B, Imai K, Shapiro JN. 2016. An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions. American Journal of Political Science **60**:783–802.

Ruppert KA et al. 2020. Use of Specialized Questioning Techniques to Detect Decline in Giraffe Meat Consumption. Journal for Nature Conservation **62**:126029.

Sagoe D, Cruyff M, Spendiff O, Chegeni R, de Hon O, Saugy M, van der Heijden PGM, Petróczi A. 2021. Functionality of the Crosswise Model for Assessing Sensitive or Transgressive Behavior: A Systematic Review and Meta-Analysis. Frontiers in Psychology **12**.

Sandbrook C, Clark D, Toivonen T, Simlai T, O’Donnell S, Cobbe J, Adams W. 2021. Principles for the socially responsible use of conservation monitoring technology and data. Conservation Science and Practice:1–12.

Solomon J, Jacobson SK, Wald KD, Gavin M. 2007. Estimating Illegal Resource Use at a Ugandan Park with the Randomized Response Technique. Human Dimensions of Wildlife **12**:75–88.

St John FAV, Keane AM, Edwards-Jones G, Jones L, Yarnell RW, Jones JPG. 2012. Identifying indicators of illegal behaviour: Carnivore killing in human-managed landscapes. Proceedings of the Royal Society B: Biological Sciences **279**:804–812.

St John FAV, Keane AM, Milner-Gulland EJ. 2013. Effective conservation depends upon understanding human behaviour. Pages 344–361 in D. W. Macdonald and K. J. Willis, editors. Key Topics in Conservation Biology 2First Edit. Oxford.

St John FAV, Linkie M, Martyr DJ, Milliyanawati B, McKay JE, Mangunjaya FM, Leader-Williams N, Struebig MJ. 2018. Intention to kill: Tolerance and illegal persecution of Sumatran tigers and sympatric species. Conservation Letters:e12451.

St John FAV, Mai C-H, Pei KJC. 2015. Evaluating deterrents of illegal behaviour in conservation: Carnivore killing in rural Taiwan. Biological Conservation **189**:86–94.

Stubbe JH, Chorus AMJ, Frank LE, de Hon O, van der Heijden PGM. 2014. Prevalence of use of performance enhancing drugs by fitness centre members. Drug Testing and Analysis **6**:434–438.

Walsh MT. 2007. Pangolins and politics in the Great Ruaha valley, Tanzania Symbol, ritual and difference. 2007.

Wilfred P, Milner-Gulland EJ, Travers H. 2019. Attitudes to illegal behaviour and conservation in western Tanzania. Oryx **53**:513–522.

Wolter F, Preisend¨orfer P. 2013. Asking sensitive questions: An evaluation of the Randomised Response Technique versus Direct Questioming using Individual validation data. Sociological Methods & Research **42**:321–353.

Yang Z. 2011. “Lucky” numbers, unlucky consumers. Journal of Socio-Economics **40**:692–699.

Yu JW, Tian GL, Tang ML. 2008. Two new models for survey sampling with sensitive characteristic: Design and analysis. Metrika **67**:251–263.