Review

Measuring pro-environmental behavioural determinants using closed-answer surveys

Hunter Doughty[1] and Laura Thomas-Walters*[2]

Abstract Changing human behaviours is a key facet of addressing global environmental issues. There are many factors (i.e. determinants) that could influence whether an individual engages in pro-environmental behaviour, and understanding these determinants can improve efforts to protect and restore the natural environment. However, despite published criticism of poor survey design, there is little practical guidance on how to capture these determinants accurately in closed-answer surveys (those with predefined answer options). A recent literature review summarized behavioural determinants of pro-environmental behaviour. We build on this by providing practical insights into how 17 key pro-environmental behavioural determinants can be measured through closed-answer surveys. We reviewed 177 papers published during 2013–2023 that met the criteria for inclusion. These papers captured 624 measurements of the 17 determinants. We found seven types of question formats used, including scales (Likert scales, semantic scales and a pictorial scale), multiple-choice questions (where respondents could select either one or more answer options), binary questions and ranking questions. We then synthesized design considerations both specifically for each format and more broadly across surveys. These considerations included using validated measures, reducing cognitive burden and biases (e.g. social desirability bias, order effects, recall bias), selecting the question format (e.g. different formats of multiple-choice or binary questions) and using best practices for scale questions. The insights collected through this review provide practical advice for developing closed-answer surveys that robustly and usefully measure key determinants of pro-environmental behaviour.

Keywords Attitudes, behaviour change, behavioural antecedents, campaigns, conservation, human dimensions, interventions, social science

The supplementary material for this article is available at doi.org/10.1017/S0030605324001194

*Corresponding author, lathowal@gmail.com
[1]Wildlife Works, Mill Valley, California, USA
[2]Yale University, New Haven, Connecticut, USA

Received 30 December 2023. Revision requested 3 April 2024. Accepted 20 August 2024. First published online 23 January 2025.

Introduction

Determinants of pro-environmental behaviour

There are widespread and increasing efforts to address human dimensions of conservation, which are now recognized as critical to achieving global environmental goals (Bennett et al., 2017; United Nations, 2023). At the individual level, pro-environmental behaviour generally refers to conservation lifestyle behaviours (e.g. household actions), social environmentalism (e.g. peer interactions and group membership), environmental citizenship (e.g. civic engagement) and land stewardship (e.g. support for conservation; Larson et al., 2015). To influence any human behaviour to be pro-environmental, or at least better for conservation, it is necessary to understand the determinants resulting in more or less of a target behaviour (Steg & de Groot, 2010; van Valkengoed et al., 2022). These determinants are grounded in an individual's perception of themselves and others, what they feel is important or good and their personal experiences, amongst other things. A recently published review provides the most comprehensive summary to date of individual behavioural determinants of pro-environmental behaviour (van Valkengoed et al., 2022). Of the 23 determinants identified (see supplementary material in van Valkengoed et al., 2022) we selected 17 that we considered most relevant to large-scale surveys, based on discussions and work over a 6-month period for Natural England, UK, which conducts annual environmentally focused surveys. We thus based our selection on a determinant's applicability to multiple pro-environmental behaviours that could be measured nationwide and its ability to be measured within the context of a larger survey without adding excessive cognitive burden. Table 1 defines these 17 determinants.

Measuring behavioural determinants

There are a multitude of ways by which pro-environmental behaviour and its determinants can be measured, such as through laboratory or field-based studies of actual behavioural decisions (e.g. energy meter readings; Lange & Dewitte, 2019). More commonly, however, we must rely on self-reported measures, whereby respondents disclose the information themselves (Newing et al., 2010). These measures can be collected through methods such as surveys, interviews or focus groups, which can be formatted

This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

Table 1 Definitions of 17 pro-environmental behavioural determinants identified by van Valkengoed et al. (2022) and assessed in this review, with the search term used for each in Google Scholar (Google, 2023b).

Behavioural determinant	Definition	Search term
Ascription of responsibility	The extent to which people personally feel responsible for the negative environmental consequences of their actions (van Valkengoed et al., 2022).	('Ascription of responsibility') AND (Environmental OR Climate)
Attitudes	The degree to which a person positively or negatively evaluates a particular behaviour (van Valkengoed et al., 2022).	(Attitudes) AND (Environmental OR Climate)
Connection to nature	The extent to which humans see themselves as part of nature (Barragan-Jason et al., 2022).	(Connected OR Connectedness OR Connection OR Relatedness) AND (Nature)
Descriptive norms	The extent to which people believe others engage in a behaviour (van Valkengoed et al., 2022). May also be referred to as 'subjective norms'.	('Descriptive norms') AND (Environmental OR Climate)
Environmental concern	Concern, worry or fear about environmental problems (van Valkengoed et al., 2022); sometimes subsumed under attitudes or knowledge.	(Concern OR Fear OR Anxiety) AND (Environmental OR Climate)
Environmental self-identity	The extent to which people think of themselves as pro-environmental (van Valkengoed et al., 2022).	('Self-identity' OR identity) AND (Environmental OR Climate)
Habits	Unconscious routines in behaviour (van Valkengoed et al., 2022).	(Habits) AND (Environmental OR Climate)
Injunctive norms	The extent to which people believe a behaviour is commonly approved or disapproved of by people or groups (van Valkengoed et al., 2022).	('Injunctive norms') AND (Environmental OR Climate)
Knowledge	Understanding of the scientific facts about the causes & impacts of environmental problems (van Valkengoed et al., 2022).	((Knowledge) AND (Environmental OR Climate)
Outcome efficacy	The extent to which people perceive their behaviour as effective at contributing to resolving environmental problems (van Valkengoed et al., 2022). May also be referred to as 'response efficacy'.	('Outcome efficacy') AND (Environmental OR Climate)
Personal norms	A person's perceived moral obligation to engage in or abstain from a particular behaviour (van Valkengoed et al., 2022). May also be referred to as 'moral norms' or 'moral obligations'.	('Personal norms') AND (Environmental OR Climate)
Problem awareness	The awareness that performing or not performing a certain behaviour increases environmental problems (van Valkengoed et al., 2022).	('Problem awareness' OR 'Awareness of consequences') AND (Environmental OR Climate)
Risk perception	An individual's evaluation of the likelihood & severity of a particular environmental hazard (van Valkengoed et al., 2022). Risk perception in this context does not refer to the perceived risk of penalties associated with not complying with a given policy or law (which is how the term is used in the context of illicit behaviours).	('Risk perception') AND (Environmental OR Climate)
Self-efficacy	The extent to which people feel capable of implementing a specific action (van Valkengoed et al., 2022). May also be referred to as 'perceived behavioural control'.	('Self-efficacy' OR 'Personal efficacy') AND (Environmental OR Climate)
Self-focused emotions	Emotions, often negative, that people feel in response to their own environmental behaviour, including guilt, shame & pride (van Valkengoed et al., 2022).	('Self-focused emotions' OR Guilt OR Pride) AND (Environmental OR Climate)
Trust	Inconsistently defined. May refer to an individual's perception that an actor (stakeholder, messenger, policy, etc.) is competent, objective, fair, reliable and caring, or that the actor will act ethically & in line with the individual's needs & wants (Wynveen & Sutton, 2015; Amin & Tarun, 2020; van Valkengoed et al., 2022).	(Trust) AND (Environmental OR Climate)
Values	Concepts that transcend specific situations/actions & serve as guiding principles for an individual in their life (Steg & de Groot, 2012).	(Values) AND (Environmental OR Climate)

using closed-answer questions with predefined answer options that the respondent must select from and/or open-answer questions that allow the respondent to answer in any way they choose. The data collected through closed-answer questions can be defined as quantitative in the sense that they are analysable using statistical methods (even if those methods are for non-numerical data; i.e. categorical data), whereas open-answer questions could

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

be considered to collect quantitative or qualitative data (i.e. data that would require transformation to be analysable statistically).

Conservationists can use any combination of social science method, question and data type, each of which has different merits and drawbacks (for a detailed overview of applied social science methods for conservation, see Newing et al., 2010). Open-ended questions, for instance, can be useful for not limiting respondents to predefined answer options, and interviews, instead of surveys, allow more extensive time and thus discourse with a respondent. However, an open-answer interview also takes more time to conduct and analyse and is more difficult to replicate. Closed-answer surveys, on the other hand, may limit some variability and depth in answers, but they are a popular method across conservation because they take less time to conduct and analyse, especially with large sample sizes, and are easier to replicate, thus making them easier to test and refine more precisely.

However, despite the prevalence of closed-answer surveys and the available guidance for conducting human research and surveys in particular (Newing et al., 2010; Sutherland et al., 2018), there has been criticism of survey robustness in conservation research (St. John et al., 2014). Given the recognized importance of addressing human dimensions of conservation, it is critical that, regardless of their background in the social sciences, conservationists have the tools to precisely and usefully measure factors influencing human behaviour (Bennett et al., 2017). We conducted this review to assess how 17 key pro-environmental behavioural determinants are being measured globally by conservationists using closed-answer surveys. We synthesize practical insights to increase the consistency, accuracy and ease of measuring pro-environmental behavioural determinants.

Methods

For this review we ran a literature search on each of the 17 pro-environmental behavioural determinants (Table 1). To be included, a study had to meet the following criteria: (1) measure one or more of the 17 pro-environmental behavioural determinants; (2) use a closed-answer survey; (3) include the text used for measuring the determinant (e.g. all of the scale statement(s) and the scale itself) and some reflective/reasoning text regarding measurement methodology; (4) relate to environmental fields; (5) be a peer-reviewed paper; (6) be primary research; (7) be published in 2013 or later (to capture the last 10 years); and (8) be published in English (this restriction was because of author capacity and we acknowledge this limitation).

We used Google Scholar (Google, 2023b), which has a higher search term character limit than many other search engines, is not limited by publisher, country or language and is an all-text search service (i.e. it looks for search terms throughout publications, not just in titles and abstracts). To reduce bias from our past search history and affiliations, we ran our searches using an incognito window in Google Chrome. Google Scholar ranks search results by relevance to search terms and by factors such as how recently and often a paper has been cited (Google, 2023a).

Prior to running the full search in February 2023 we conducted multiple test searches using variations of the search terms. We used four a priori-identified papers recommended by experts in the field to test the search structure. We found three of the four papers to be identifiable directly in the search results and one to be identifiable indirectly, as the author had other similar papers displayed in the search results. We performed 17 separate searches, one on each determinant (Table 1). We then combined all 17 search terms with 'AND (Survey OR Closed-answer OR Questionnaire OR Poll OR Measure)' to capture papers with survey methods, as well as 'AND (Behavio* OR Nudg*)' to capture behaviour-related papers.

We screened the top 30 hits per determinant. We found this number produced a large amount of high-quality data (almost all hits warranted a detailed assessment for inclusion) and resulted in information that began to repeat itself, indicating thematic saturation in the search results. We extracted the following information from each paper: research details (e.g. year and country), whether the paper tested a behavioural model, theory or paradigm, overall survey methodology, pro-environmental behaviour measurement details (if behaviour and not behaviour intent was measured) and measurement details of behavioural determinants (e.g. question format and number of questions).

Literature review findings

We screened 510 papers and included 177 published during 2013–2023. These papers captured 624 measurements of the 17 behavioural determinants, covering 48 countries or country combinations. Some 58% of measurements were offline, 44% were online (some were both online and offline) and 69% of measurements were done in the context of a theory explicitly discussed by the authors. The full data are available in Supplementary Material 1, and findings on behavioural theories discussed by the authors are available in Supplementary Material 2.

Question formats

Seven types of question formats were used in the literature to assess the 17 behavioural determinants and pro-environmental behaviour (if it was measured; Table 2). These formats included scales (Likert scales, semantic scales and pictorial scales), multiple-choice questions

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

TABLE 2 Results from the literature review in terms of question formats used in closed-answer surveys to assess pro-environmental behavioural determinants (Table 1) and whether these determinants were assessed via a single or multiple questions in a given study. The last row indicates whether the question formats were also used to directly measure pro-environmental behaviour in these papers.

Behavioural determinant	Likert scale	Binary	Multiple choice: 1+ answer	Multiple choice: 1 answer	Ranking	Semantic scale	Pictorial scale	Single question	Multiple questions
Ascription of responsibility	X		X		X			X	X
Attitudes	X	X				X		X	X
Connection to nature	X		X				X	X	X
Descriptive norms	X							X	X
Environmental concern	X	X						X	X
Environmental self-identity	X	X							X
Habits	X							X	X
Injunctive norms	X		X					X	X
Knowledge	X	X		X				X	X
Outcome efficacy	X							X	X
Personal norms	X							X	X
Problem awareness	X			X				X	X
Risk perception	X							X	X
Self-efficacy	X							X	X
Self-focused emotions	X	X	X					X	X
Trust	X							X	X
Values	X							X	X
Pro-environmental behaviour	X	X		X					X

(where respondents could select one or more answer options), binary questions and ranking questions.

Scale questions The most common question format was scales. For this review we define scale questions as those whereby respondents rate statements in standalone questions (but it should be noted that the term 'scale' can also be used to describe an overarching prescribed set of questions that are used in combination, such as the New Ecological Paradigm scale; Stern et al., 1995).

In standalone Likert-scale questions, respondents are presented with one or multiple statement items and then asked to rate their agreement with or strength of feeling for each item (Wang et al., 2016). The phrasing of these answer scales varied, for example: 'Strongly disagree–Strongly agree', 'Never–Always', or 'Not at all important–Extremely important'. We found Likert scales to be used for every determinant and pro-environmental behaviour.

We also found the semantic differential scale (i.e. bipolar scale) used to assess attitudes. In a semantic scale the statement item often references a behaviour, situation or policy, amongst others, and respondents may be asked to rate this item multiple times across different scales (the sum of these ratings is seen as a single attitudinal measure). For example, Liu et al. (2017) asked respondents to rate their attitude towards car transport reduction along four seven-point scales: 'Harmful–Beneficial', 'Disgusting–Pleasant', 'Bad–Good', and 'Unworthy–Valuable'.

We found one pictorial scale (i.e. visual scale) used. The Inclusion of Nature in Self scale assesses the connection of respondents to nature (Schultz, 2002), in which respondents are shown Venn diagrams in which the two circles represent themselves and nature. This seven-point scale has seven images of circles that vary in how much they overlap, ranging from completely separate to fully overlapping (Liefländer et al., 2013).

There is substantial literature on the use of scales and what constitutes best practice (Boateng et al., 2018; Jebb et al., 2021). Important considerations for scale items include using single- versus multi-item measurements, reverse coding and item order. Eighty-nine per cent of studies using scale questions included multiple item measurements for a given determinant. Using multiple items decreases the probability that any one item will skew results and permits assessments of internal consistency as a basis for factor analysis (see the Design considerations section below). Furthermore, as most items are positively or negatively framed, authors should include items framed from opposing value orientations (such items are then reverse coded). Reverse-coded items reduce social desirability bias (see

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

Conclusion section) and increase the probability of capturing the true perspective of a respondent, as they should theoretically answer opposingly on such items. Regarding item order, when including reverse-coded items it is useful to mix the order of statements so that not all positive/pro-environmental statements come first or last. Similarly, all measured determinants should be randomized, or at least mixed, to avoid order effects (see Conclusion section; Lacroix & Gifford, 2018) and reduce the probability of respondents confounding their perception of one item with another similar item measuring the same construct (Pakpour et al., 2014).

Scale orientation is also important. Although some authors employed a positive-to-negative scale, such as 'agree' to 'disagree', the greater tendency was to use a negative-to-positive scale, such as 'disagree' to 'agree'. We posit that the latter should be preferred because starting with a negative option may help lessen priming effects for statements that are often pro-environmentally framed.

In addition, the number of points along the answer scale varied from four to 10, with the most common being five-point and seven-point scales. There has been extensive debate regarding which of these scales is better, but seven-point scales may be preferable for measuring attitude-like constructs as they reduce the psychological distance between points on the scale and provide more granularity in the data for analysis without overwhelming respondents with too large a scale (Wakita et al., 2012; Joshi et al., 2015). Additionally, despite the proliferation of odd-numbered scales, having a midpoint/neutral option may not always be best. Taufique et al. (2017, p. 9) purposefully used a four-point scale to encourage respondents 'to choose a positive or negative response to minimise social desirability bias', and because 'the omission of a midpoint is particularly useful when dealing with Asian respondents, who often have a higher mid-range response tendency'. See Chyung et al. (2017) for an often-cited resource on determining whether to use a scale midpoint, considering factors such as whether a midpoint would increase response rate whilst still maintaining data quality.

Scales are widely used because of the nuance they provide, but they can also be cognitively burdensome to respondents (McLeod et al., 2011). As such, it is important to consider participant fatigue across a survey and whether/when a scale format is best.

Multiple-choice questions Multiple-choice questions were used to assess injunctive norms, knowledge, problem awareness, ascription of responsibility, connection to nature, self-focused emotions and pro-environmental behaviour. Authors used this format most when measuring knowledge and pro-environmental behaviour. Respondents to multiple-choice questions were either able to select a single answer option (i.e. mutually exclusive answers) or multiple answer options (i.e. non-mutually exclusive answers; Libarkin et al., 2018; Zhu et al., 2020).

Although non-mutually exclusive answers were only seen in three studies, binary questions in other studies could have been reformatted into this type of multiple-choice question. For example, Vesely & Klöckner (2018) asked respondents 56 separate yes/no questions regarding their past pro-environmental behaviours. These questions could be merged into one multiple-choice question where respondents select any of the 56 behaviours. A benefit of asking many yes/no questions is that doing so may encourage respondents to think specifically about each behaviour. However, a single multiple-choice question probably reduces cognitive burden and potentially provides more accurate answers as respondents can select fewer behaviours without feeling the potential guilt of answering many questions with a 'no' or switching back and forth between 'yes' and 'no' (which relates to the internal desire of respondents to feel consistent in their behaviours; Vesely & Klöckner, 2018).

As with scales, the order of answer options should be considered when developing multiple-choice questions. Primacy and recency effects, as part of the serial position effect, can cause respondents to focus on the first and last answer options in a list (Murdock Jr, 1962). In addition, when looking for a correct answer, as is the case when measuring respondent knowledge, respondents also tend to look to the middle answer option, particularly if they are unsure (Attali & Bar-Hillel, 2003). Randomizing answers helps reduce these biases.

Binary questions Questions with binary answer options (yes/no or true/false) were used to assess attitudes, knowledge, environmental concern, self-focused emotions, environmental self-identity and pro-environmental behaviour. Similar to multiple-choice questions, authors used this format the most when measuring knowledge and pro-environmental behaviour. Binary questions are exemplified in Roczen et al. (2014), where respondent attitudes were measured using both Likert-scale questions and 23 yes/no items, such as 'I get up early to watch the sunrise'.

A major consideration for binary questions is whether to include an 'I don't know' or 'Prefer not to say' option. In 7 of 14 studies to use binary questions, the authors did include this option. Both Bolderdijk et al. (2013) and Ünal et al. (2018) reasoned that incorporating an 'I don't know' option in their true/false questions would mean that respondents were not forced to guess the right answer when they didn't know it, thereby enabling the authors to assess respondent knowledge more accurately. Additionally, when assessing self-focused emotions (e.g. guilt), a third answer option gives respondents the ability to opt out of answering instead of forcing them to inaccurately label themselves if they do not know (Hickman et al., 2021). However, the usefulness of this opt-out option depends on the aim of the

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

study. For instance, if study authors want to encourage respondents to make a choice or state whether they perform a behaviour (especially when the answer is relatively straightforward, such as whether the respondent regularly gets up early to watch the sunrise), then having an opt-out answer could reduce the usable data points. Data from this opt-out option are often treated as missing (although sometimes they are grouped with 'no/false'), thereby decreasing the number of respondents with a completed survey that the authors can use in analyses, which in turn decreases the statistical power to detect effects.

Ranking questions We found one ranking question used. Zeng et al. (2020) assessed ascription of responsibility by first asking respondents 'Who should take the responsibility for environmental protection?' via a multiple-choice question with non-mutually exclusive answers. Then they asked respondents to rank their selected answers in order of who they think is most responsible for environmental protection (e.g. (1) Government, (2) Every individual, (3) Business enterprises, and (4) Others). This provided a creative closed-answer approach to gain nuance from respondent answers without needing to employ an open-ended question.

Similar to the other question formats, the order in which answer options are presented can influence the ranking order given by respondents (Serenko & Bontis, 2013), and as such it is important to randomize answer options. Additionally, surveyors should consider how the respondent will physically create their ranking to minimize cognitive burden. Blasius (2012) found that in web surveys a drag-and-drop user interface performed better than a numbering, arrows or most–least interface at increasing substantive answers and reducing dropout and non-response rates.

Design considerations

A number of biases could affect the outcome and accuracy of a closed-answer survey. Some sources of data error may be non-directional (i.e. errors across respondents balance each other out if the sample size is large enough), but this is difficult to ascertain pre-emptively, so it is best to consider all potential errors as biases to be mitigated where possible.

One major bias to consider is non-response bias, which refers to gaps in data on the behaviours and perceptions of the individuals who do not participate in either the whole survey or in answering specific questions, making the data non-representative of the population (Davern, 2013). This bias could be mitigated through survey design (e.g. incentivising respondents) and during analysis (e.g. weighting data to match the population; Okafor, 2010). Incomplete surveys or half-hearted answers can also result from a survey placing too much cognitive burden on a respondent, causing them to lose interest or become overwhelmed (i.e. cognitive fatigue); thus survey length, clarity and question ease are also important considerations.

Any self-reported answer is also inherently subject to respondent perspectives, memories and intentions to convey a certain image of themselves (Althubaiti, 2016). All reported behaviour is prone to recall bias: humans have faulty memories and often recall their own behaviour inaccurately even when attempting to be accurate (Althubaiti, 2016). Measures of pro-environmental behaviour are especially prone to this bias (Koller et al., 2023). Tactics such as asking respondents to recall short timeframes or prompting their recall by using memorable temporal landmarks (e.g. national holidays) are helpful (Gaskell et al., 2000).

A key driver of self-reporting biases in surveys is social desirability bias (Wheeler et al., 2019). This occurs when respondents consciously or subconsciously modify their responses to match what they think the surveyor wants to hear. It results from the inherent tendency of humans to want to appear socially desirable and to maintain a positive self-image (Latkin et al., 2017). To help mitigate this bias, respondents should be informed about the anonymity and confidentiality of their responses and that there are no right or wrong answers (Esfandiar et al., 2020). The way questions are phrased can also greatly affect this bias and thus needs careful consideration. Leading questions (e.g. 'Do you agree that wiping out all animals on the planet is a bad thing?') probably induce this bias, but more subtle factors influence it as well. Leviston & Uren (2020), for example, discuss how loosely specified behaviours (e.g. changing one's gardening practices) are more prone to social desirability bias than concrete behaviours (e.g. installing a rainwater tank or insulation). In addition, although behavioural determinants have been most commonly assessed via direct questions, if the topic is particularly sensitive to respondents then any direct question, no matter how carefully crafted, may produce biased results. As such, conservationists should consider whether proxy measurements or specially designed indirect questions (i.e. sensitive questioning techniques) are more appropriate (Nuno & St. John, 2015; Cerri et al., 2021).

Similarly, priming and order can influence respondent answers. Priming occurs when the respondent is prompted to think about a certain topic or identification with a certain group before answering a question, which could be unintentional on the part of the surveyor (Hjortskov, 2017). For instance, if surveyors ask questions about the child of a respondent and then ask questions about the respondent (e.g. their personal norms), the respondent might now be primed to think about their child and answer the follow-up questions with a greater focus on the legacy impacts of their behaviour on future generations. Thus, the sequence of questions throughout a survey, and the order of statements within a scale or of answer options within a multiple-choice,

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

binary and ranking question can all affect what a respondent is thinking about and how they think they should answer a given question (Lacroix & Gifford, 2018).

Considering such biases, it is crucial to carefully design how the overall survey is presented to respondents, as well as how each question and answer is phrased and ordered. There is a wealth of advice available on this. For example, Bruine de Bruin (2011) assesses framing effects on survey questions, and Althubaiti (2016) considers response biases such as recall and social desirability bias, and how to mitigate these effects. For developing scale questions, Jebb et al. (2021) and Boateng et al. (2018) provide advice specific to Likert scales (with many principles relating to other question formats). Conservationists can also consider, and test, whether forcing a response will decrease non-response bias and social desirability bias whilst not increasing half-hearted/non-substantive answers; this has been discussed earlier for midpoint scales and opt-out answers added to binary questions (Chyung et al., 2017; Ünal et al., 2018) but is applicable to all question types.

We identified three tactics that conservationists used to develop survey questions. In 76% of measurements the authors relied on previous research (e.g. scales previously validated by other authors) to design their questions, in 28% of measurements the authors piloted/pre-tested their survey questions and in 14% of measurements the authors used a panel of experts to design their questions. Employing all three tactics is arguably best practice. For instance, Pagiaslis & Krontalis (2014) used existing literature to develop a survey draft that was reviewed by a panel of five experts in consumer research and biofuels. The resulting questionnaire was then piloted with 150 consumers before the final survey was conducted. If translation work was necessary in a study, authors often used multiple additional steps to ensure the survey conveyed the same concepts in the other language and culture. Nguyen et al. (2016), for example, used a prescribed back-translation technique involving two professional translators in English and Vietnamese, followed by a review from two other bilingual researchers and then an expert panel review and in-depth consumer interviews. Niamir et al. (2020) provide another example of using all three question development tactics and translation steps. It is also recommended to check that scales have comparable psychometric properties after translation, such as through differential item functioning (Petersen et al., 2003).

It is particularly important to contextualize surveys when respondents are children. Surveyors must consider and test whether their question-and-answer options are relevant to younger respondents and how these respondents will interpret them. Nine studies involved respondents under 18 years old, and although most authors probably considered their audience, only Wallis & Loy (2021, p. 5) explained survey adjustments made for these respondents: 'Based on studies with adolescents and young people. . .we asked for social influences in the form of the perceived pro-environmental activism of their parents and friends.' There were, however, some adult-focused studies that adapted surveys for audiences with different literacy levels. For example, Farage et al. (2021, p. 4) stated: 'based on our participants' background (e.g. literacy level, less practice in expressing opinions and making distinctions)' they used a four-point scale represented visually as four circles of varying colour. The answer options were also written on each circle (i.e. dark green with 'Strong agreement' written; light green with 'Agreement' written; dark red with 'Strong disagreement' written; light red with 'Disagreement' written). Respondents could then point to the circle they wished to select.

Lastly, validation is key across all measurement constructs (e.g. scales) in a survey. To validate constructs, surveyors test the validity of a survey (i.e. whether the survey measures what it is intended to measure; Tsang et al., 2017) and reliability. Validity can take many forms, including but not limited to criterion validity (i.e. how well scores on the survey correlate with relevant external, non-test criteria) and construct validity (e.g. whether the scale measures the construct of focus, itself indicated by convergent validity, discriminant validity, differentiation by known groups and correlation analysis; Boateng et al., 2018). Similarly, reliability can take multiple forms, such as test–retest consistency (i.e. whether the survey would give the same results if it was repeated with the same people) and internal consistency (e.g. whether all the items in the scale measure the same variable consistently, often measured with Dillon-Goldstein's rho or using the split-half reliability coefficient; Robinson, 2018; Revelle & Condon, 2019). The steps taken to validate a survey vary across fields, but they can involve tactics such as expert panels, piloting, testing–retesting the same respondents and statistical analyses (Tsang et al., 2017). Given that validation helps ensure surveys measure what the surveyor intended, conservationists ought to go beyond the three-pronged survey development tactic discussed earlier to confirm that newly developed constructs are reliable and valid. There is a wealth of literature on how to validate constructs such as psychometric scales (e.g. Boateng et al., 2018; Hughes, 2018).

Summary and application

Key considerations

Using these insights into question formats and design considerations, we now discuss key recommendations for choosing question types and designing closed-answer surveys. Conservationists should base any decisions they make regarding these recommendations on their specific study context and audience, taking into account factors

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

such as audience age, status in a household, literacy, cultural, financial and/or religious upbringing, and wider socio-political trends, sensitivities, physical environment and access.

Use validated measures To ensure surveys reliably measure what is intended, use validated measures where possible, but do not assume a question that has been validated elsewhere will necessarily work in a new study context: validation is still necessary. Alternatively, develop surveys through existing literature, expert panels and piloting, as well as through any further steps needed for validation. Translation of surveys requires additional steps.

Select appropriate question formats To increase the comprehension and ease of respondents, consider which question format is best for the specific information and audience (e.g. children), such as a scale, multiple-choice, binary or ranking format. Additionally, consider whether to use multiple-choice questions with mutually exclusive or non-mutually exclusive answers, binary questions with or without a third neutral/opt-out answer, or separate binary questions, or one multiple-choice question with non-mutually exclusive answers (Table 3, Example 2).

Use best practices for scale questions To increase the accuracy and usefulness of scales, use best-practice guidance such as multiple items with opposing value orientations per each determinant, scales with a seven-point range starting from the negative, selecting the right scale (e.g. semantic) and range (e.g. Never–Always) and considering whether alternative question formats would reduce cognitive burden (Table 3, Example 1).

Mitigate non-response bias and survey fatigue To increase response rate and quality of participation, consider factors such as incentives, weighting respondents and survey length and understandability. Additionally, consider whether to remove opt-out/neutral answer options.

Mitigate social desirability bias To reduce the influencing of respondent results, reassure respondents that answers are anonymous, use insights on phrasing and order (such

Table 3 Hypothetical question formats seen in conservation research, along with potential alternatives created using insights discussed in this review.

Common question format	Potential alternative(s)	
Example 1		
How much do you agree or disagree with the following statement? (a) I am concerned that there are fewer animals in the forest. *Select one answer.* (1) Strongly agree (2) Agree (3) Neither agree nor disagree (4) Disagree (5) Strongly disagree	How much do you disagree or agree with the following statements? *[Randomized and interspersed amongst other statements]* (a) I am concerned that there are fewer animals in the forest. (b) I am not really worried that there are fewer animals in the forest. *Select one answer per statement.* (1) Strongly disagree (2) Disagree (3) Somewhat disagree (4) Neither disagree nor agree (5) Somewhat disagree (6) Agree (7) Strongly agree	*Select one answer* What statement do you agree with the most regarding whether there are fewer animals in the forest? (a) I am in no way concerned (b) I am not concerned (d) I am concerned (e) I am very concerned
Example 2		
Select one answer per question. Do you eat bushmeat? (a) Yes (b) No Do you eat chicken? (a) Yes (b) No Do you eat fish? (a) Yes (b) No Do you eat beef? (a) Yes (b) No	*Select all answers that apply.* Which of the following do you eat on a weekly basis? (a) Chicken (b) Fish (c) Bushmeat (d) Beef	

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

as asking about concrete behaviours), consider sensitive questioning techniques and consider the potential removal of neutral answer options.

Mitigate priming and order effects To reduce the influencing of respondent results, randomize question and answer item order when possible or at least consider how earlier questions influence later questions and biases such as serial position effects within answer options.

Mitigate against recall bias To increase the accuracy of self-reported behaviours, use tactics such as asking about short timeframes, using temporal landmarks and asking about concrete behaviours.

See Supplementary Material 3 for possible measurement approaches for pro-environmental behaviour and each determinant based on the formats we found to be most commonly employed and validated in the literature.

Applying considerations

To increase the usefulness and application of this review for conservationists, we have used the insights discussed above to modify two hypothetical questions (Table 3).

In Example 1, first alternative, we changed the scale to a seven-point scale and rearranged it to start with 'Strongly disagree'. We rephrased the question and included two statement items with opposing value orientations, varying the wording slightly to ensure results are not the result of a lack of understanding of the phrasing (Lacasse, 2016; that both statements are perceived in relatively the same way by the audience should be tested during the piloting phase). Ideally, the order of these statement items would also be randomized and interspersed amongst other scale items in the survey.

In Example 1, second alternative, we changed the scale to a multiple-choice question with mutually exclusive answers. This alternative requires a single question for participants to answer instead of two scale statements to rate and thus could reduce cognitive burden. The question is phrased neutrally and the answer options are ordered with the least environmentally desirable option first. Given the inherent ordered nature of these answer options, for participant comprehension we kept them in order instead of randomizing their order (as would normally be recommended).

In Example 2 we converted the four binary questions into a single multiple-choice question with non-mutually exclusive answers to reduce cognitive burden. We rephrased the question to increase the specificity and to narrow the timescale to one that is probably easier for respondents to recall. Additionally, we placed the presumed answer of interest (bushmeat) as a middle option to reduce primacy and recency effects. Because this question is not asking for a correct answer, we were, however, less concerned about the middle option being a bias. In this example we are assuming that eating bushmeat is not highly sensitive, otherwise an indirect questioning technique may be more appropriate.

Conclusion

Addressing pro-environmental behaviour is critical to achieving global conservation aims, and influencing any behaviour often requires understanding its underlying drivers. Through this literature review we assessed how 17 key determinants of pro-environmental behaviour are commonly measured using closed-answer surveys. Given that these determinants span a range of topics that are important to furthering conservation, such as human attitudes, norms and values, we believe that the guidance presented here will be relevant across conservation globally. We have synthesized practical insights, from using validated measures to addressing recall and social desirability biases, to support conservationists in designing surveys more easily, robustly and consistently.

Acknowledgements We thank the UK Government's Department for Environment, Food and Rural Affairs (Defra) for their funding and support of this research, and Ruth Lamont of Defra for her support and feedback.

Author contributions Study design, data collection: both authors; data analysis: HLD; writing: HLD; revision: both authors.

Conflicts of interest None.

Ethical standards No ethical approval was required for this research. The research abided by the *Oryx* guidelines on ethical standards.

Data availability The full data are available in Supplementary Material 1.

References

Althubaiti, A. (2016) Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of Multidisciplinary Healthcare*, 9, 211–217.

Amin, S. & Tarun, M.T. (2020) Effect of consumption values on customers' green purchase intention: a mediating role of green trust. *Social Responsibility Journal*, 17, 1320–1336.

Attali, Y. & Bar-Hillel, M. (2003) Guess where: the position of correct answers in multiple-choice test items as a psychometric variable. *Journal of Educational Measurement*, 40, 109–128.

Barragan-Jason, G., de Mazancourt, C., Parmesan, C., Singer, M.C. & Loreau, M. (2022) Human–nature connectedness as a pathway to sustainability: a global meta-analysis. *Conservation Letters*, 15, e12852.

Bennett, N.J., Roth, R., Klain, S.C., Chan, K., Christie, P., Clark, D.A. et al. (2017) Conservation social science: understanding and integrating human dimensions to improve conservation. *Biological Conservation*, 205, 93–108.

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

Blasius, J. (2012) Comparing ranking techniques in web surveys. *Field Methods*, 24, 382–398.
Boateng, G.O., Neilands, T.B., Frongillo, E.A., Melgar-Quiñonez, H.R. & Young, S.L. (2018) Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health*, 6, 149.
Bolderdijk, J.W., Gorsira, M., Keizer, K. & Steg, L. (2013) Values determine the (in)effectiveness of informational interventions in promoting pro-environmental behavior. *PLOS One*, 8, e83911.
Bruine de Bruin, W. (2011) Framing effects in survey design: how respondents make sense of the questions we ask. In *Perspectives on Framing* (ed. G. Keren), pp. 303–324. Taylor & Francis, London, UK.
Cerri, J., Davis, E.O., Verissimo, 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.
Chyung, S.Y., Roberts, K., Swanson, I. & Hankinson, A. (2017) Evidence-based survey design: the use of a midpoint on the Likert scale. *Performance Improvement*, 56, 15–23.
Davern, M. (2013) Nonresponse rates are a problematic indicator of nonresponse bias in survey research. *Health Services Research*, 48, 905–912.
Esfandiar, K., Dowling, R., Pearce, J. & Goh, E. (2020) Personal norms and the adoption of pro-environmental binning behaviour in national parks: an integrated structural model approach. *Journal of Sustainable Tourism*, 28, 10–32.
Farage, L., Uhl-Haedicke, I. & Hansen, N. (2021) Problem awareness does not predict littering: a field study on littering in the Gambia. *Journal of Environmental Psychology*, 77, 101686.
Gaskell, G.D., Wright, D.B. & O'Muircheartaigh, C.A. (2000) Telescoping of landmark events: implications for survey research. *The Public Opinion Quarterly*, 64, 77–89.
Google (2023a) *About Google Scholar*. Google, Mountain View, USA. scholar.google.com/intl/en/scholar/about.html [accessed 14 March 2023].
Google (2023b) *Google Scholar*. Google, Mountain View, USA. scholar.google.co.uk [accessed 1 February 2023].
Hickman, C., Marks, E., Pihkala, P., Clayton, S., Lewandowski, R.E., Mayall, E.E. et al. (2021) Climate anxiety in children and young people and their beliefs about government responses to climate change: a global survey. *The Lancet Planetary Health*, 5, e863–e873.
Hjortskov, M. (2017) Priming and context effects in citizen satisfaction surveys. *Public Administration*, 95, 912–926.
Hughes, D.J. (2018) Psychometric validity: establishing the accuracy and appropriateness of psychometric measures. In *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development* (eds P. Irwing, T. Booth & D.J. Hughes), pp. 751–779. Wiley Blackwell, Hoboken, USA.
Jebb, A.T., Ng, V. & Tay, L. (2021) A review of key Likert scale development advances: 1995–2019. *Frontiers in Psychology*, 12, 637547.
Joshi, A., Kale, S., Chandel, S. & Pal, D.K. (2015) Likert scale: explored and explained. *Current Journal of Applied Science and Technology*, 7, 396–403.
Koller, K., Pankowska, P.K. & Brick, C. (2023) Identifying bias in self-reported pro-environmental behavior. *Current Research in Ecological and Social Psychology*, 4, 100087.
Lacasse, K. (2016) Don't be satisfied, identify! Strengthening positive spillover by connecting pro-environmental behaviors to an 'environmentalist' label. *Journal of Environmental Psychology*, 48, 149–158.
Lacroix, K. & Gifford, R. (2018) Psychological barriers to energy conservation behavior: the role of worldviews and climate change risk perception. *Environment and Behavior*, 50, 749–780.
Lange, F. & Dewitte, S. (2019) Measuring pro-environmental behavior: review and recommendations. *Journal of Environmental Psychology*, 63, 92–100.
Larson, L.R., Stedman, R.C., Cooper, C.B. & Decker, D.J. (2015) Understanding the multi-dimensional structure of pro-environmental behavior. *Journal of Environmental Psychology*, 43, 112–124.
Latkin, C.A., Edwards, C., Davey-Rothwell, M.A. & Tobin, K.E. (2017) The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. *Addictive Behaviors*, 73, 133–136.
Leviston, Z. & Uren, H.V. (2020) Overestimating one's 'green' behavior: better-than-average bias may function to reduce perceived personal threat from climate change. *Journal of Social Issues*, 76, 70–85.
Libarkin, J.C., Gold, A.U., Harris, S.E., McNeal, K.S. & Bowles, R.P. (2018) A new, valid measure of climate change understanding: associations with risk perception. *Climatic Change*, 150, 403–416.
Liefländer, A.K., Fröhlich, G., Bogner, F.X. & Schultz, P.W. (2013) Promoting connectedness with nature through environmental education. *Environmental Education Research*, 19, 370–384.
Liu, Y., Sheng, H., Mundorf, N., Redding, C. & Ye, Y. (2017) Integrating norm activation model and theory of planned behavior to understand sustainable transport behavior: evidence from China. *International Journal of Environmental Research and Public Health*, 14, 1593.
McLeod, A., Pippin, S. & Wong, J.A. (2011) Revisiting the Likert scale: can the fast form approach improve survey research? *International Journal of Behavioural Accounting and Finance*, 2, 310–327.
Murdock Jr, B.B. (1962) The serial position effect of free recall. *Journal of Experimental Psychology*, 64, 482–488.
Newing, H., Eagle, C.M., Puri, R. & Watson, C.W. (2010) *Conducting Research in Conservation: Social Science Methods and Practice*. Routledge, Abingdon, UK.
Nguyen, T.N., Lobo, A. & Greenland, S. (2016) Pro-environmental purchase behaviour: the role of consumers' biospheric values. *Journal of Retailing and Consumer Services*, 33, 98–108.
Niamir, L., Ivanova, O., Filatova, T., Voinov, A. & Bressers, H. (2020) Demand-side solutions for climate mitigation: bottom-up drivers of household energy behavior change in the Netherlands and Spain. *Energy Research & Social Science*, 62, 101356.
Nuno, A. & St. John, F.A.V. (2015) How to ask sensitive questions in conservation: a review of specialized questioning techniques. *Biological Conservation*, 189, 5–15.
Okafor, F. (2010) Addressing the problem of non-response and response bias. *CBN Journal of Applied Statistics*, 1, 7.
Pagiaslis, A. & Krontalis, A.K. (2014) Green consumption behavior antecedents: environmental concern, knowledge, and beliefs. *Psychology & Marketing*, 31, 335–348.
Pakpour, A.H., Zeidi, I.M., Emamjomeh, M.M., Asefzadeh, S. & Pearson, H. (2014) Household waste behaviours among a community sample in Iran: an application of the theory of planned behaviour. *Waste Management*, 34, 980–986.
Petersen, M.A., Groenvold, M., Bjorner, J.B., Aaronson, N., Conroy, T., Cull, A. et al. (2003) Use of differential item functioning analysis to assess the equivalence of translations of a questionnaire. *Quality of Life Research*, 12, 373–385.
Revelle, W. & Condon, D.M. (2019) Reliability from α to ω: a tutorial. *Psychological Assessment*, 31, 1395–1411.
Robinson, M.A. (2018) Using multi-item psychometric scales for research and practice in human resource management. *Human Resource Management*, 57, 739–750.

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press

Roczen, N., Kaiser, F.G., Bogner, F.X. & Wilson, M. (2014) A competence model for environmental education. *Environment and Behavior*, 46, 972–992.

Schultz, P.W. (2002) Inclusion with nature: the psychology of human-nature relations. In *Psychology of Sustainable Development* (eds P. Schmuck & W.P. Schultz), pp. 61–78. Springer, Boston, USA.

Serenko, A. & Bontis, N. (2013) First in, best dressed: the presence of order-effect bias in journal ranking surveys. *Journal of Informetrics*, 7, 138–144.

Steg, L., & de Groot, J. (2010) Explaining prosocial intentions: testing causal relationships in the norm activation model. *British Journal of Social Psychology*, 49, 725–743.

Steg, L. & de Groot, J. (2012) Environmental values. In *The Oxford Handbook of Environmental and Conservation Psychology* (ed. S.D. Clayton), pp. 81–92. Oxford University Press, Oxford, UK.

Stern, P.C., Dietz, T. & Guagnano, G.A. (1995) The new ecological paradigm in social-psychological context. *Environment and Behavior*, 27, 723–743.

St. John, F.A.V., Keane, A.M., Jones, J.P.G. & Milner-Gulland, E.J. (2014) FORUM: robust study design is as important on the social as it is on the ecological side of applied ecological research. *Journal of Applied Ecology*, 51, 1479–1485.

Sutherland, W.J., Dicks, L.V., Everard, M. & Geneletti, D. (2018) Qualitative methods for ecologists and conservation scientists. *Methods in Ecology and Evolution*, 9, 7–9.

Taufique, K.M.R., Vocino, A. & Polonsky, M.J. (2017) The influence of eco-label knowledge and trust on pro-environmental consumer behaviour in an emerging market. *Journal of Strategic Marketing*, 25, 511–529.

Tsang, S., Royse, C.F. & Terkawi, A.S. (2017) Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anaesthesia*, 11, S80–S89.

Ünal, A.B., Steg, L. & Gorsira, M. (2018) Values versus environmental knowledge as triggers of a process of activation of personal norms for eco-driving. *Environment and Behavior*, 50, 1092–1118.

United Nations (2023) *The 17 Goals*. Department of Economic and Social Affairs, United Nations, New York, USA. sdgs.un.org/goals [accessed 23 August 2023].

van Valkengoed, A.M., Abrahamse, W. & Steg, L. (2022) To select effective interventions for pro-environmental behaviour change, we need to consider determinants of behaviour. *Nature Human Behaviour*, 6, 1482–1492.

Vesely, S. & Klöckner, C.A. (2018) Global social norms and environmental behavior. *Environment and Behavior*, 50, 247–272.

Wakita, T., Ueshima, N. & Noguchi, H. (2012) Psychological distance between categories in the Likert scale: comparing different numbers of options. *Educational and Psychological Measurement*, 72, 533–546.

Wallis, H. & Loy, L.S. (2021) What drives pro-environmental activism of young people? A survey study on the Fridays for Future movement. *Journal of Environmental Psychology*, 74, 101581.

Wang, S., Fan, J., Zhao, D., Yang, S. & Fu, Y. (2016) Predicting consumers' intention to adopt hybrid electric vehicles: using an extended version of the theory of planned behavior model. *Transportation*, 43, 123–143.

Wheeler, S.A., Gregg, D. & Singh, M. (2019) Understanding the role of social desirability bias and environmental attitudes and behaviour on South Australians' stated purchase of organic foods. *Food Quality and Preference*, 74, 125–134.

Wynveen, C.J. & Sutton, S.G. (2015) Engaging the public in climate change-related pro-environmental behaviors to protect coral reefs: the role of public trust in the management agency. *Marine Policy*, 53, 131–140.

Zeng, J., Jiang, M. & Yuan, M. (2020) Environmental risk perception, risk culture, and pro-environmental behavior. *International Journal of Environmental Research and Public Health*, 17, 1750.

Zhu, W., Yao, N., Guo, Q. & Wang, F. (2020) Public risk perception and willingness to mitigate climate change: city smog as an example. *Environmental Geochemistry and Health*, 42, 881–893.

https://doi.org/10.1017/S0030605324001194 Published online by Cambridge University Press