ORIGINAL RESEARCH ARTICLE

Finding quality in quantitative methods: The case of food forestry in the Netherlands

Anna M. Roodhof[1] 🔘 | **Esther J. Veen**[2] 🔘

[1] Rural Sociology Group, Wageningen Univ., Hollandseweg 1, Wageningen 6706KN, the Netherlands

[2] Aeres Univ. of Applied Sciences, Arboretum West 98, Almere 1325WB, the Netherlands

Correspondence
Esther J. Veen, Aeres Univ. of Applied Sciences, Arboretum West 98, Almere 1325WB, the Netherlands.
Email: e.veen@aeres.nl

Assigned to Associate Editor Debra Davidson.

Abstract

In current nature valuation efforts, sociocultural values are often overlooked. Calls for the importance of including these values often advise to use a mixed-method approach; however, restrictions concerning budget and time often prevent researchers from doing so. In this study we offer a compromise—a questionnaire that aims to capture sociocultural values of nature as closely as possible. It was created as part of the Dutch National Monitoring Programme Food Forests (*Nationaal Monitoringsprogramma Voedselbos*sen [NMVB]), hence pertaining to food forestry (FF) in the Netherlands. We used three strategies. First, we reviewed literature on nature valuation frameworks to establish the constructs for the questionnaire, finding that the survey needs to be sensitive to a plurality of worldviews. Second, we consulted survey methodology to accommodate for this plurality and to minimize nonresponse and measurement error. Third, we tested the questionnaire extensively. Our final questionnaire combines closed- and open-ended questions to examine sociocultural values from different angles to allow participants to elaborate on previously given answers in their own words and to captivate participants, prompting continued attentiveness. We believe that the questionnaire-creation process can be inspiring to other researchers. Moreover, we call upon them to use our questionnaire and to continuously evaluate and revise it in the process.

1 | INTRODUCTION

Assessing sociocultural values—the importance people assign to objects, processes, or phenomena—is an exceptionally complex task, yet many scholars have emphasized their noteworthiness in nature valuation efforts (Bieling et al., 2014; Fagerholm et al., 2016; Pascual et al., 2017; Scholte et al., 2015; Tress & Tress, 2001). Anything can be a sociocultural value if people appreciate that specific element.

However, values differ from person to person and are subject to change. They are not self-explanatory; values require interpretation, which in turn entails a thorough understanding of the cultural context (Stephenson, 2008).

Scholars often promote mixed-method approaches to investigate values (Arias-Arévalo et al., 2018; Chan et al., 2012a; Scholte et al., 2015) using questionnaires to complement qualitative approaches or vice versa (Plieninger et al., 2013). However, budgetary constraints often lead to a reliance on quantitative approaches alone (Dillman et al., 2014). Hence, rather than stressing the importance of moving beyond such approaches, in this paper we argue for an optimization of

---

Abbreviations: ES, ecosystem services; FF, food forestry; NMVB, *Nationaal Monitoringsprogramma Voedselbossen*.

quantitative approaches. Building on the case study of measuring sociocultural values of food forests in the Netherlands, we demonstrate three strategies that helped us capture the multiplicity and qualitative nature of these values when using a questionnaire.

Our first strategy was a thorough literature review on existing nature valuation frameworks to create the constructs to be measured by our questionnaire. The review shows that as valuation frameworks are rooted in particular worldviews, research instruments often limit participants in their ability to express ideas and perceptions of nature outside of that particular worldview. We argue, therefore, that more explicit attention needs to be given to a variety of frameworks and their underlying worldviews. Our second strategy was to carefully consider the current literature on survey methodology, finding ways to accommodate for this plurality of worldviews as well as aiming to minimize measurement error and nonresponse. The third strategy consisted of extensively testing and adjusting the questionnaire. We elaborate on each of these strategies in a separate section of this paper and end with a concluding section in which we reflect upon the strategies. First, however, we shortly present our case study—food forestry (FF) in the Netherlands.

2 | CASE STUDY: FOOD FORESTS IN THE NETHERLANDS

FF is a form of agroforestry. While agroforestry encompasses any land-use system "in which trees are grown in association with agricultural crops, pastures, or livestock" (Young, 1997, p. 3), FF can be defined as a land-use system that mimics the ecosystem of a natural forest predominantly using edible and perennial plant species (Park et al., 2018). FF has origins in tropical home gardens, which Kumar and Nair (2004) have labelled as "the epitome of sustainability" (p. 135), as these gardens typically contain a wide array of plant species (Kumar & Nair, 2004). Likewise, FF systems have a relatively high biodiversity compared with conventional agricultural systems (Park et al., 2018). This gives ground for beneficial plant–plant interactions, which implies that much like a natural forest, food forests require little management by the farmer. Moreover, food forests render the use of external inputs (such as chemical fertilizers and pesticides) redundant (Albrecht & Wiek, 2020). Besides food production, food forests provide a habitat for varying wildlife species, and they sequester more carbon than conventional farms (Park & Higgs, 2018). As such, FF encompasses both food production and nature conservation (Green Deal Voedselbossen, 2017a). It could therefore play a role in both food security and environmental regeneration (Albrecht & Wiek, 2020; Wartman et al., 2018). However, while research efforts are increasingly directed toward illuminating the ecological and socioeconomic benefits of

---

Core Ideas

- Sociocultural values can be incorporated in quantitative nature valuation efforts.
- It is important to include different frameworks and worldviews in such attempts.
- A questionnaire can be designed to mimic qualitative approaches to some degree.
- Extensive testing helps creating better questionnaires.
- We urge others to use our instrument and evaluate its efficacy in the process.

---

food forests in the global north, sociocultural values have yet to be studied (Park & Higgs, 2018; Wartman et al., 2018).

The Dutch National Monitoring Programme Food Forests (*Nationaal Monitoringsprogramma Voedselbossen* [NMVB]) continuously measures the developments of food forests in the Netherlands. Recently, NMVB sought to expand their quantitative measuring tool to include sociocultural valuation of FF. The questionnaire-creation process described in this paper was conducted as part of a Master of Science project of the first author and supervised by the second author. During the creation of the questionnaire—especially with regards to the literature review of nature valuation frameworks and the execution of the testing strategies—the first author worked in close collaboration with Suzy Rebisz, another Master of Science student from Wageningen University whose project centered on the outcomes of the questionnaire.

3 | UNDERSTANDING SOCIOCULTURAL VALUES

Several researchers have tried to study values of nature and green, such as those of agroforestry (Fagerholm et al., 2016), ecosystem services (ES) (Scholte et al., 2015), landscape (Tress & Tress, 2001; Bieling et al., 2014), and nature more broadly (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, 2016; Pascual et al., 2017). In this section we first review how researchers have tried to measure values of nature to reflect upon the use of different valuation frameworks. We end by addressing how we applied the findings from these frameworks to our questionnaire.

3.1 | Ecosystems services framework

As the need for nature valuation frameworks at large is predominantly based on their prospective capacity to inform policy makers about trade-offs that occur when altering

ecosystems (Chan et al., 2012a), valuation frameworks often take the ES framework as a starting point. This framework identifies quantifiable aspects of nature to calculate favorable trade-offs, which is facilitated by an understanding of nature in terms of services. Payments for ES, for instance, allow policymakers to compensate for the loss of ES when a trade-off has been made. Hence, the ES framework bridges ecological and economic interests in nature; it allows policymakers to put a price on ecological functions and thereby to calculate how to compensate for ecological losses (Chan et al., 2012a). In other words, the ES framework implicitly understands value as a measurable unit (Pascual et al., 2017).

This implies, however, that when trying to value ES, values tend to be wrongfully conflated with services and benefits (Chan et al., 2012b). Chan et al. (2012b) therefore separate these concepts, defining benefits as "valued goods or experiences" (p. 10), which can be generated by services, the ecosystem processes that underpin benefits. Values refer to "the preferences, principles and virtues that we (up)hold as individuals or groups" (p. 10), where "*principles* generally pertain to characteristics of an action or decision (the means), *preferences* pertain to the consequences of an action (the ends), and *virtues* pertain to the actor(s)" (Chan et al., 2012b, p. 11, emphasis added). People's judgment of nature can thus be guided by underlying principles based on personal beliefs of what is right and wrong, people's desires about the world, or what kind of person they want to be (i.e., what kind of virtues they seek to uphold). As such, nature valuation efforts need to address more than simple measurement units.

Indeed, scholars have promoted several improvements to the ES framework, for instance by including nonmaterial ES such as the aesthetic value of nature (Chan et al., 2012b). However, the language used in this framework renders it difficult to depart from an understanding of nature as serving people as implied in the word service. An understanding of the ES concept that includes different visions of human engagement with nature would therefore be more fitting (Arias-Arévalo et al., 2018).

3.2 | A plurality of worldviews

Some frameworks depart from ES as the main object of evaluation; those by Tress and Tress (2001) and Stephenson (2008) offer insightful approaches to studying human–nature interactions. By offering a more holistic definition of landscape, they recommend a far broader understanding of nature valuation. Tress and Tress (2001) propose a systems approach to studying the interaction between people and the landscapes they inhabit and rooted in the premise that while landscapes set the conditions for human existence, people also actively change the landscape in return. Accordingly, the authors denote five dimensions of landscapes (spatial, mental, temporal, the land-

scape as nexus of nature and culture, and the landscape as a complex system). Similarly, Stephenson (2008) emphasizes the cultural importance of landscapes, stating that while certain landscapes might be perceived as ordinary by most people, they can hold significant value for those in proximity to these landscapes. Stephenson (2008) defines landscape as consisting of the three components: forms, relationships, and practices. Hence, unlike the ES framework, which approaches the value of nature only in terms of its utility to people (Muradian & Pascual, 2018), Tress and Tress (2001) and Stephenson (2008) stress the importance of sociocultural values of nature, which cannot be reduced to a single value indicator (Arias-Arévalo et al., 2018).

The concept of worldviews also stresses the need to go beyond a singular understanding of nature. Worldviews disclose how societies or cultures position themselves in relation to nature (Blaser, 2013). Muradian and Pascual (2018) operationalized worldviews as relational models that characterize the relation between people and nature (see Table 1). These relational models capture the specific aspect of worldviews that drives sociocultural value attribution. Nature valuation efforts are often rooted in one particular worldview, because the researchers, or valuing agents, behind it are rooted in this respective worldview (Arias-Arévalo et al., 2018). In other words, how researchers study phenomena and interpret the results is affected by their worldview. To secure transparent, valid, and reliable research, researchers therefore need to disclose their worldview and reflect upon it. Such a reflexive approach to bias facilitates the documentation of the worldviews of research participants, which may differ from those of the researchers. In sum, worldviews are intricately connected to how people attribute value to nature, which people do in different ways. It is therefore essential that different worldviews are considered in nature valuation efforts (Arias-Arévalo et al., 2018).

As nature valuation in terms of ES refers (only) to the "provision of direct and indirect benefits to people from ecosystems" (Chan et al., 2012b, p. 8), this framework excludes worldviews that diverge from this perspective on human–nature relationships (Arias-Arévalo et al., 2018). More inclusive frameworks exist, but these generally call for an extensive qualitative approach to capture the plurality of worldviews. While we agree that a plurality of worldviews should be considered, our aim is to try and do so by using a questionnaire.

3.3 | Measuring worldviews in food forestry

In the creation of our questionnaire, we have been cautious of using words that assume a certain worldview, as a 'singular valuation language' can result in participants' inability to answer the question accurately: either the question itself or the

T A B L E 1 A typology of human–nature relational models adopted from Muradian and Pascual (2018)

Relational model	View on the human–nature interaction
Detachment	Nature is not important. It needs to make space for technology
Domination	Humans are entitled to control and destroy nature
Devotion	Nature is sacred and humans need to make sacrifices to show it respect
Stewardship	Humans are a part of nature and need to care for nature to preserve it
Wardship	Nature is a separate entity to be protected from humans
Ritualized exchange	Nature and humans are equal and co-exist in search of an equitable balance
Utilization	Nature is a resource to be extracted for human needs or pleasure

response alternatives may not capture their underlying perception of human–nature interactions (Arias-Arévalo et al., 2018). For example, we followed Chan et al. (2012b) and Scholte et al. (2015) in refraining from conflating the terms 'value' and 'service', for these two concepts are only interchangeable in the utilization relational model. Additionally, the questionnaire offers participants a moment to reflect on their answer after each question using open-ended questions. This enables them to give more specific answers in their own words, which is necessary to ensure that participants can share their values as accurately as possible (Scholte et al., 2015).

To include a plurality of worldviews in our questionnaire, we used Muradian and Pacual's (2018) conceptualization of worldviews as well as the existing literature on valuation frameworks to establish four constructs: (a) perception of the human–nature relationship, (b) perception of the food forest, (c) attitude toward the food forest, and (d) personal well-being.

The first construct is based on Muradian and Pascual's (2018) typology of human–nature relational models. The second construct forms a bridge between worldview and values, as it enquires about how participants perceive their environment. This is affected by participants' worldviews, and in turn, it affects how they value their environments. The actual sociocultural value of FF is measured using the third construct. Finally, existing valuation frameworks indicated the importance of perceived health benefits in nature valuation efforts (Bieling et al., 2014), so we created a fourth construct to address this. We operationalized these four constructs into multiple variables and corresponding indicators as well as several variables to address potential confounding factors (including basic demographic indicators as age, level of education, and income).

3.3.1 | Perception of the human–nature relationship

We started the questionnaire by addressing how participants understand the human–nature relationship, because participants' worldviews affect how they attribute value to FF (Arias-Arévalo et al., 2018). We derived statements from each

human–nature relational model as described in Table 1. However, as both 'detachment' and 'utilization' view nature as separate from humans, either as unimportant or as a resource to be exploited, we merged them (Muradian & Pascual, 2018). We also merged 'devotion' and 'wardship' because both models view nature as a separate entity with intrinsic rights. Consequently, we included five questions to explore participants' views on human–nature relationships.

3.3.2 | Perception of the food forest

People's direct perceptions of the landscape affect their attitudes toward that landscape and how they interact with it (Tress & Tress, 2001). We therefore included questions that investigate how participants perceive the food forest itself. To connect values to specific parts of the food forest, we invited participants to upload a photo of the food forest, upon which we asked them why they chose that image and what it represented. Moreover, we used a free-listing technique (asking participants to list all the words they associate with a concept) (Russell Bernard, 2011) to investigate how people perceive the concept of FF.

3.3.3 | Attitude toward the food forest

We also created questions based on the conceptualization of values by Chan et al. (2012b). These authors differentiate between individual and community values. The questionnaire addresses both types of values by asking participants how they think food forests contribute to them personally and to society. It enquires about contributions to society at different scales: local, regional, national, and global. Additionally, we included three closed-ended questions in which participants' attitudes toward the food forest are addressed. One question concerned participants' reasons for being involved with the food forest, the other two questioned which aspects of the food forest they consider most important. All are nominal questions: in the first one, participants are asked to select a minimum of one and a maximum of three reasons for being

involved with the food forest, in the second they are asked to select which aspect they deem most important, and in the third, participants are offered the same response alternatives, but asked to select which aspects they also find important. The response alternatives are based on types of values as identified by the value typology of Chan et al. (2012b). This inclusive value typology distinguishes between principles, preferences, and virtues, as explained above, but also between market-mediated vs. nonmarket-mediated values, self-oriented and other-oriented values, individual and group values, experiential and metaphysical values, transformative and non-transformative values, and anthropocentric and biocentric values.

3.3.4 | Personal well-being

Finally, we aimed to explore how food forests contribute to participants' well-being, because personal well-being and sociocultural value attribution are closely related (Bieling et al., 2014). It is difficult to objectively pinpoint how food forests might contribute to well-being as participants might be happy and healthy because of unrelated circumstances—the list of potential confounding factors is extensive. As such, we followed Bieling et al. (2014), who suggest simply asking participants how they think the landscape contributes to their personal well-being; perceived linkages between landscape and well-being are an appropriate indicator to measure this.

In sum, the existing literature helped us operationalize the concept of worldviews and values into the four constructs described above and clarified that it is important to incorporate a plurality of worldviews. It was also informative regarding which factors are important to consider in nature valuation, such as individual vs. collective values or values pertaining to different scales. However, this body of literature did not demonstrate how to ask questions that lead to accurate answers. To measure the four constructs accurately, we therefore used the existing literature on survey methodology as elaborated upon in the next section.

4 | REDUCING METHODOLOGICAL ISSUES

As argued in the introduction, literature on nature valuation frameworks emphasizes the importance of qualitative research methods, but, in reality, such methods are not always feasible. Hence, although we acknowledge that a singular, quantitative method might not fully capture all dimensions of values participants attribute to food forests, we contend that it is useful to try to capture sociocultural values by using a questionnaire. After understanding and operationalizing the

constructs we aimed to measure, we carefully considered methodological issues to find ways to create a strong quantitative tool. We constructed our questions and arranged them in a way to (a) include a plurality of worldviews and (b) minimize nonresponse and measurement error, which is a common problem in (online) questionnaires. While we discuss them separately, these issues are heavily related, as both contribute to our primary aim of collecting authentic, elaborate participant responses.

4.1 | Including a plurality of worldviews

To include a plurality of worldviews we approached sociocultural values from different angles. Moreover, we used different types of questions, especially combinations of open- and closed-ended questions. In face-to-face interviews, researchers can use probing techniques when participants offer incomplete answers. Direct probing is not possible in a questionnaire. We therefore employed follow-up questions to stimulate participants to think deeper about their answers and provide more precise and accurate data. By adding open-ended questions to closed-ended ones, we enabled participants to address the nuances of their answers in the preceding closed-ended question. In this way, we aimed to obtain in-depth answers and mitigate the lack of complexity in closed-ended questions; however, we kept the open-ended questions optional. While this can lead to higher nonresponse rates for these questions, requiring participants to answer them can cause them to quit the questionnaire altogether, increasing the overall nonresponse rate.

We considered several factors when crafting closed-ended questions. Firstly, closed-ended questions that assess opinions or attitudes can use bipolar or unipolar scales. The Likert scale is an often-used bipolar scale. It measures both direction (i.e., agreement or disagreement) and intensity (i.e., completely agree or somewhat agree), requiring respondents to complete a complex mental task. Likert scales should therefore only be used when they serve a clear purpose (Russell Bernard, 2011). As unipolar questions only assess intensity, they yield more reliable results (Alwin et al., 2018). We mainly employed unipolar scales for ordinal closed-ended questions but chose a Likert scale when enquiring about the extent to which respondents agree with statements about human–nature interactions. In the latter case, we wanted to document both direction and intensity about participants' beliefs. For ordinal questions, we limited the number of response options to five; a higher number of options decreases reliability (Alwin et al., 2018). For bipolar questions, we offered four response options so that participants cannot choose a middle category, though we included a 'don't know' option. For unipolar questions this was not needed; respondents can simply select the zero option (i.e., not well).

Secondly, as reliability declines when the number of response alternatives increases in categorical questions (Alwin et al., 2018), we had to make a trade-off between reliability and precision. In some cases, we were able to limit the number of response alternatives to five. In other cases, the questions required additional response alternatives to ensure participants could answer the question accurately. In case the response alternatives did not reflect participants' preferred answers, we provided the option 'other'.

In short, we used a combination of open- and closed-ended questions to document participants' sociocultural values of FF in different formats. The closed-ended questions, which result in quantifiable data, are at risk of measurement error, as the predetermined response alternatives might not reflect participants' experiences. We therefore carefully considered how to structure and phrase them. Additionally, we combined them with open-ended questions, which result in qualitative data. These questions ensure that participants can elaborate on the selected response alternative in the preceding question in a more unfiltered way. How the use of different question formats relates to measurement error is the topic of the next section.

4.2 | Minimizing nonresponse and measurement error

Self-administered questionnaires—we focus on a digitally self-administered survey in this paper, as this administration mode was set by the NMVB—are prone to nonresponse and measurement error (Dillman et al., 2014). We therefore gave these issues particular attention, heavily relying on the work of Dillman et al. (2014) and Gobo and Mauceri (2014). Nonresponse error occurs when a reduced response rate skews the outcome. A prevalent cause of nonresponse error is the questionnaire and communication mode, which can both incentivize and disincentivize respondents to participate (Goldberg & Sciarini, 2019; Jäckle et al., 2010; Mavletova & Couper, 2014; Vannieuwenhuyze et al., 2010). Measurement error occurs when respondents give inaccurate answers. This can be due to a myriad of factors, which Gobo and Mauceri (2014) summarize as intentional errors, errors of memory, and errors of communication. Specifically, measurement error can occur when respondents are unable or unwilling to answer the questions accurately (Dillman et al., 2014), when they do not recall behavior or perceptions correctly or when they are not aware of motivations underpinning their behavior (Gobo & Mauceri, 2014). Measurement error can be systematic across all respondents or vary between respondents (Dillman et al., 2014).

An important way to minimize such error is to create trust between researcher and respondent; when there is little trust, the response rate might suffer, or participants may be unmotivated to answer accurately. A self-administered questionnaire can be seen as a social interaction in which the questionnaire functions as a proxy for the interviewer with which the respondent interacts (Bradburn, 2016; Dillman et al., 2014). We actively sought to establish trust by employing an extensive invitation strategy. Mixing invitation and reminder modes can result in an overall higher response rate (Mavletova & Couper, 2014). We sent out personalized emails in which we introduced ourselves and invited participants to call us. Over the phone, we were able to establish rapport. It was only then that we sent out the invitations to the actual questionnaire.

We also aimed to decrease nonresponse by increasing participation benefits. We clearly communicated the value of the questionnaire and its results to participants. Moreover, we included interesting questions in the design (Dillman et al., 2014) and strategically ordered questions to make the answering process more engaging (Dillman et al., 2014; Stern et al., 2007). For example, we included the previously described free-listing association question and the photo question. To minimize cognitive efforts required from participants—reducing participant costs—we provided an indication of the number of words participants were expected to list. Moreover, alternating between closed- and open-ended questions also helps to make a questionnaire more interesting. Hence, using different types of questions was not only intended to provoke participants to think about their opinions in different ways but also to retain their attention throughout the questionnaire. It was also a way to establish rapport with participants before asking boring or more intrusive, personal questions such as those about demographic characteristics.

Another way to minimize nonresponse is to minimize questionnaire length, which is important especially for web and mail questionnaires (Peytchev et al., 2020). We limited our questionnaire to 53 questions, which was expected to be answerable in 20 minutes. We focused on the most relevant dimensions of values rather than trying to capture many dimensions. We grouped questions in three sections, and to avoid long batteries of questions, which are particularly prone to question order effects (Dillman et al., 2014), we included filtering questions that enable asking follow-up questions only to relevant respondents.

As lexical effects, such as ambiguous words, also play a considerable role in the occurrence of measurement error (Gobo & Mauceri, 2014), we paid specific attention to wording, avoiding terms that can be interpreted differently as well as 'loaded words' (i.e., words that invoke positive or negative emotions). Likewise, question order affects how questions are interpreted; while a specific sequence of questions can facilitate understanding, it may also lead to bias. However, randomizing question order can confuse respondents (Dillman et al., 2014). We therefore grouped questions of the same topic together to create context while keeping in mind that participants are influenced by each question preceding the next.

In sum, we included different types of questions to foster trust, minimize participant cost, and increase participant
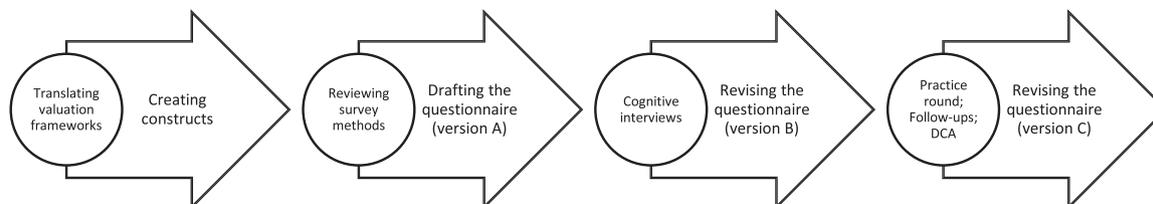
FIGURE 1   The iterative process of creating the questionnaire

benefits. In this way, we aimed to minimize nonresponse and measurement error, while the variety of question types, especially the combination of closed and open-ended questions, was meant to help people consider a plurality of worldviews. Nevertheless, the resulting questionnaire was based on literature alone. In practice, participants might encounter unanticipated issues such as interpretation issues. We therefore took an additional step to ensure the efficacy of our instrument; we conducted several rounds of testing. We explain these different rounds and their results in the following section.

5 | TESTING

After creating a first draft of the questionnaire (Version A), we conducted a preliminary round of testing in the form of cognitive interviews, which led to an updated version of the questionnaire (Version B). We then conducted three more types of testing—a practice round, follow-up interviews, and a deviant case analysis—to ensure the quality of the research instrument. This resulted in the final version of the questionnaire (Version C), which we included as supplemental material. Our testing strategies are graphically represented in Figure 1.

5.1 | Testing Version A: Cognitive interviews

We used cognitive interviews to assess the quality of questionnaire Version A. These interviews followed the 'think aloud' pretesting strategy listed by Gobo and Mauceri (2014). Participants were asked to read everything they would normally read out loud when completing a questionnaire and verbalize their thought process when answering the questions. This enabled us to see the questionnaire through the eyes of the participants (Gobo & Mauceri, 2014) and to perceive it separately from the theoretical underpinnings on which it was based. Because of time constraints, we only conducted two cognitive interviews with personal connections who are involved with food forests. The interviewees were a man in his early 30s and a woman in her early 20s, both with a Bachelor of Science degree. This sample was not meant to be representative of the target population, as the purpose of this testing strategy was to serve as a rough preliminary check of the questionnaire.

We recommend a larger and more diversified sample when possible.

Despite this small sample, the interviews provided a useful first check of the questionnaire, which led to several useful insights. We used these to improve the questionnaire into Version B. Participants of the cognitive interviews were able to answer most of the questions without issues and interpreted the vast share of questions as expected. However, the interviews shed light on five points of improvement: confusion about section headings, confusion about question order, confusion about question structure, inability to answer questions, and interpretative issues (for more detail, see Supplemental Table S1). We made changes to the questionnaire for each of these five issues. Some of these changes were simple (e.g., adjusting the wording), but in other cases, changes required more structural adjustments such as adding or omitting questions. Overall, questions in the adjusted version of the questionnaire are more direct and precise to ensure participants' interpretations are similar, and response alternatives more fully capture participants' answers.

5.2 | Testing Version B: Practice round, follow-up interviews, and deviant case analysis

We used three testing strategies to test Version B of the questionnaire. First, we sent it to a first group of respondents. This so-called practice round allowed us to detect general issues with the software used and the format of the questionnaire as well as to examine participants' answers to open questions. The research sample selected for the practice round was a nonprobability sample. It consisted of persons affiliated with the current 25 participating food forests of NMVB. We had obtained contact information of one or two representatives per food forest (most often owner, manager, or volunteer) and asked them to share our questionnaire within their teams. Ultimately, 49 responses from 23 food forests were recorded. Table 2 provides some characteristics of the respondents. Because of the nonprobability sample and the fact that both the distribution of food forests and how participants were related to the food forests was skewed, we did not analyze the data statistically for error. We did, however, compare participants' answers to the open-ended questions with their answers to the closed-ended questions preceding them. This gave us

TABLE 2 Descriptive statistics of the participants of the practice round

Descriptive variable	Frequency
	% (*n*)
Gender	
Male	61.2 (30)
Female	36.7 (18)
Don't want to say	2 (1)
Level of education	
Post-graduate degree	8.2 (4)
MSc degree	40.8 (20)
BSc degree	42.9 (21)
Vocational education	6.1 (3)
High school diploma	2 (1)
Age	
<30	20 (9)
30–40	13 (6)
41–50	24 (11)
51–60	26 (12)
>60	17 (8)
Employment status	
Full-time employment	12.2 (6)
Part-time employment	36.7 (18)
Freelance employment	30.6 (15)
Unemployed	20.4 (10)

TABLE 3 Descriptive statistics of the participants of the follow-up interviews

Descriptive variable	Frequency
	% (*n*)
Gender	
Male	50 (4)
Female	50 (4)
Don't want to say	0 (0)
Level of education	
MSc degree	37.5 (3)
BSc degree	50 (4)
Vocational education	12.5 (1)
Age	
<30	37.5 (3)
41–50	25 (2)
51–60	37.5 (3)
Employment status	
Full-time employment	25 (2)
Part-time employment	37.5 (3)
Freelance employment	12.5 (1)
Unemployed	25 (2)

an indication about the efficacy of the closed-ended questions, for example those pertaining to the response alternatives we provided. Moreover, by examining the summary statistics, we were able to identify deviant cases, discussed later in this section.

Second, we used follow-up interviews. Like cognitive interviews, follow-up interviews are meta-interviews to examine the quality of a questionnaire. However, unlike cognitive interviews, follow-up interviews are reflexive; they require participants to reflect on their response process in retrospect as opposed to sharing their thoughts during this process. Interview participants were first given the opportunity to express any general concerns or comments. We then inquired about their interpretation of each question and their ability to answer the questions within the answer formats (e.g., open or closed-ended). Although the interviews were structured by the questionnaire itself, we customized verbal probing for each participant. Some participants were very talkative and willing to share their response process, while others found reporting on this process difficult. The latter category required more extensive probing while refraining from steering too much. Out of the 49 participants of the practice round, 31 people indicated their willingness to participate in a follow-up interview. We invited eight participants from different food forests for

an interview, based on their demographic characteristics, to ensure a diversified sample, see Table 3.

The practice round and the follow-up interviews indicated that, to a large extent, participants interpreted questions correctly and offered elaborate answers in the open questions. As such, we gained ample insight in the sociocultural value of FF. Inevitably, however, several issues occurred. We identified two types of issues: (a) structural issues, such as limitations of the software used and (b) those that could be resolved by adjusting the questionnaire or the communication strategy. While the first could not be solved within the timeframe of this study, the second led to several changes in the questionnaire. There were issues pertaining to participants' inability to answer the question, lack of depth in participants' answers, interpretative issues, redundant questions, and confusion about the questions or the questionnaire in general. Subsequently, we rectified unintended mistakes, provided participants with more elaborate instructions, changed the wording of questions, added or changed response categories, and omitted questions (for more detail, see Supplemental Table S2). Overall, changes were intended to improve participants' ability to understand what is asked of them and to answer accordingly so that they are better able to give an authentic answer within the format of the questionnaire.

The third strategy to test Version B of the questionnaire was a deviant case analysis, which is a technique to improve the predictive capacity of a questionnaire (Gobo & Mauceri, 2014) and to indicate interpretation issues or flaws. We

purposely invited participants for the follow-up interviews who had recorded deviant answers, which we asked participants to explain. Three deviant cases were found. The first case involved the privacy statement. As this is not a relevant topic for this paper, we do not discuss it further here. The second deviant case pertains to the question in which participants were asked whether their appreciation of food forests had changed since they first encountered FF. We expected participants to answer that their appreciation had remained the same or had improved as in informal conversations prior to the practice round participants displayed great enthusiasm about FF. However, two participants stated that their appreciation had decreased. One of these was willing to participate in a follow-up interview in which they explained that after prolonged involvement, they realized that FF is not a perfect alternative to conventional agriculture. Although this answer deviated from expectations, it affirmed the inclusion of the respective question. We rephrased the open-ended question to ask for this explanation and underlying motivation more directly. The third deviant case involved the role participants have within the food forest. One of the participants of the questionnaire indicated that their role was a 'visitor', but we knew them to be employed by a food forest. An interview with them clarified that they had misinterpreted the question. Overall, the interviews clarified that certain words draw more attention than others so that when participants read questions quickly, they may overlook certain other words, causing misinterpretation. As such, we paid specific attention to the wording of questions prone to misinterpretation.

Based on the outcome of the practice round, the follow-up interviews, and the deviant case analysis, we improved the questionnaire to create a final Version C. This is the version included as supplemental material. The extensive testing rounds showed that there is no such thing as the perfect questionnaire. We did find that as participants have different preferences and perceptions, it is important to offer room for nuance via open-ended questions. While it might be tempting to exclusively use closed-ended questions, as these are more easily analyzed statistically, the open-ended questions proved essential to capture participants' values, which tend to be complex and nuanced. Moreover, we noticed that each testing round is merely a snapshot of how participants experienced the questionnaire. The improved version, Version C, should therefore by no means be viewed as a flawless, static end-product. We recommend that each use of the questionnaire is seen as an opportunity to improve the questionnaire for subsequent use.

6 | CONCLUSION

Scholars have stressed the importance of using mixed methods when valuing nature, highlighting that deliberative approaches, such as focus groups, are essential to capture individual and group values. Time or budgetary restraints often render this impossible. A continuing reliance on qualitative methods to study sociocultural values therefore risks omitting these types of values from nature valuation efforts altogether. In fact, in current ES valuation frameworks, the omission of sociocultural values is common practice. As such, the objective of our research was to create a quantitative research instrument that measures these sociocultural values as precisely as possible, using FF in the Netherlands as a case study.

Our questionnaire is based on extensive research on both nature valuation and survey methodology. First, we scrutinized the term 'sociocultural value' and why such values are worth studying. This strategy helped us define our social constructs. We found that values are more than simply a unit of measurement used to calculate favorable trade-offs between nature and the economy (Chan et al., 2012b). While such an understanding of values certainly simplifies compensating for loss of ES, it overlooks the sociocultural significance of the landscape (Stephenson, 2008). Literature research showed that sociocultural values are rooted in corresponding worldviews, and that treating values as isolated opinions overlooks the influence of these worldviews. We therefore prioritized the inclusion of a plurality of worldviews in our questionnaire.

Second, we considered methodological issues to ensure that the questionnaire is sensitive to this plurality of worldviews and to minimize nonresponse and measurement error. These issues are highly related, as measurement error is a serious risk when questionnaires seek to capture complex social constructs, as in our work. We tackled this issue by using different types of questions, most importantly closed-ended questions followed by open-ended questions. While open-ended questions are more difficult to analyze, they allow participants to describe their perceptions and opinions in an unfiltered way. Such descriptions are key to understanding sociocultural values of nature. The combination of open- and closed-ended questions not only helped us include a variety of worldviews, but it also documents answers in different ways, leading to more precise results, and keeps participants engaged, decreasing risk of nonresponse and measurement error. Reviewing methodological literature also helped us consider the effects of question order, question format, and choice of vocabulary as well as to improve our communication strategy. We therefore stress the importance of considering survey methodology when creating an online questionnaire or a questionnaire aimed to measure complex data such as sociocultural values.

Finally, to ensure its quality, our questionnaire has been thoroughly tested using various qualitative testing strategies. This showed us that while the first two strategies—reviewing existing nature valuation frameworks and reviewing literature on survey methodology—served as a useful preparation for drafting the questionnaire, such preparation alone is not

enough. Participants encounter an array of issues answering questions: they experience interpretative issues, lack prerequisite knowledge, or are simply unable to answer the question accurately within the given question format. Several issues arose during the testing process, which we were able to address by adjusting the questionnaire. This strategy has thus proven vital to the questionnaire creating process.

While our questionnaire focuses on measuring sociocultural effects of food forests, certain parts of it will be useful for researchers studying other landscapes or land uses (please note that our final questionnaire is available as supplemental material). Moreover, our three strategies to create the questionnaire may inspire others in their questionnaire-creation processes. Although our approach to crafting the questionnaire itself is not necessarily innovative, it is exceptionally extensive, which ultimately benefits the reliability of the research instrument. We acknowledge, however, that no questionnaire can be completely without flaws and that it cannot entirely be a one-size-fits-all tool. Each use of the questionnaire should therefore be seen as an opportunity to evaluate its efficacy and its applicability to specific circumstances and contexts and to adapt it accordingly.

AUTHOR CONTRIBUTIONS
Anna M. Roodhof: Conceptualization; Data curation; Formal analysis; Methodology; Project administration; Visualization; Writing – original draft; Writing – review & editing. Esther J. Veen: Supervision; Writing – original draft; Writing – review & editing.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ORCID
Anna M. Roodhof ⓘ https://orcid.org/0000-0002-0543-0953
Esther J. Veen ⓘ https://orcid.org/0000-0002-0436-0359

REFERENCES

Albrecht, S., & Wiek, A. (2020). Food forests: Their services and sustainability. *Journal of Agriculture, Food Systems, and Community Development*, *10*, 91–105. https://doi.org/10.5304/jafscd.2021.103.014

Alwin, D. F., Baumgartner, E. M., & Beattie, B. A. (2018). Number of response categories and reliability in attitude measurement. *Journal of Survey Statistics and Methodology*, *6*, 212–239. https://doi.org/10.1093/jssam/smx025

Arias-Arévalo, P., Gómez-Baggethun, E., Martín-López, B., & Pérez-Rincón, M. (2018). Widening the evaluative space for ecosystem services: A taxonomy of plural values and valuation methods. *Environmental Values*, *27*, 29–53. https://doi.org/10.3197/096327118İ15144698637513

Bieling, C., Plieninger, T., Pirker, H., & Vogl, C. R. (2014). Linkages between landscapes and human well-being: An empirical exploration with short interviews. *Ecological Economics*, *105*, 19–30. https://doi.org/10.1016/j.ecolecon.2014.05.013

Blaser, M. (2013). Notes towards a political ontology of 'environmental' conflicts. In L. Green (Ed.), *Contested ecologies: Dialogues in the South on nature and knowledge* (pp. 13–27). Human Sciences Research Council.

Bradburn, N. M. (2016). Surveys as social interactions. *Journal of Survey Statistics and Methodology*, *4*, 94–109. https://doi.org/10.1093/jssam/smv037

Chan, K. M. A., Guerry, A. D., Balvanera, P., Sarah, K., Satterfield, T., Basurto, X., Bostrom, A., Chuenpagdee, R., Gould, R., Halpern, B. S., Hannahs, N., Levine, J., Norton, B., Ruckelshaus, M., Russell, R., Tam, J., & Woodside, U. (2012a). Where are cultural and social in ecosystem services? A framework for constructive engagement. *Bioscience*, *62*, 744–756. https://doi.org/10.1525/bio.2012.62.8.7

Chan, K. M. A., Satterfield, T., & Goldstein, J. (2012b). Rethinking ecosystem services to better address and navigate cultural values. *Ecological Economics*, *74*, 8–18. https://doi.org/10.1016/j.ecolecon.2011.11.011

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored method* (4th ed.). Wiley.

Fagerholm, N., Oteros-Rozas, E., Raymond, C. M., Torralba, M., Moreno, G., & Plieninger, T. (2016). Assessing linkages between ecosystem services, land-use and well-being in an agroforestry landscape using public participation GIS. *Applied Geography*, *74*, 30–46. https://doi.org/10.1016/j.apgeog.2016.06.007

Gobo, G., & Mauceri, S. (2014). *Constructing survey data: An interactional approach*. Sage Publications.

Goldberg, A. C., & Sciarini, P. (2019). Who gets lost, and what difference does it make? Mixed modes, nonresponse follow-up surveys and the estimation of turnout. *Journal of Survey Statistics and Methodology*, *7*, 520–544. https://doi.org/10.1093/jssam/smy025

Green Deal Voedselbossen. (2017a). *C-219 Green deal voedselbossen* (In Dutch). https://www.greendeals.nl/sites/default/files/downloads/GD219-dealtekst-Voedselbossen.pdf

Green Deal Voedselbossen. (2017b). *Nationaal monitoringsprogramma voedselbossen (NMVB)*.(In Dutch). https://www.greendealvoedselbossen.nl/nationaal-monitoringsprogramma-voedselbossen-nmvb/

Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. (2016). *Preliminary guide regarding diverse conceptualization of multiple values of nature and its benefits,*

including biodiversity and ecosystem functions and services. Plenary of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services–Fourth session, Kuala Lumpur, 22–28 February 2016. https://www.ipbes.net/sites/default/files/downloads/IPBES-4-INF-13_EN.pdf

Jäckle, A., Roberts, C., & Lynn, P. (2010). Assessing the effect of data collection mode on measurement. *International Statistics Review*, *78*, 3–20. https://doi.org/10.1111/j.1751-5823.2010.00102.x

Kumar, B. M., & Nair, P. K. R. (2004). The enigma of tropical homegardens. *Agroforestry Systems*, *61*, 135–152.

Mavletova, A., & Couper, M. P. (2014). Mobile web survey design: Scrolling versus paging, SMS versus e-mail invitations. *Journal of Survey Statistics and Methodology*, *2*, 498–518. https://doi.org/10.1093/jssam/smu015

Muradian, R., & Pascual, U. (2018). A typology of elementary forms of human–nature relations: A contribution to the valuation debate. *Current Opinion in Environmental Sustainability*, *35*, 8–14. https://doi.org/10.1016/j.cosust.2018.10.014

Park, H., & Higgs, E. (2018). A criteria and indicators monitoring framework for food forestry embedded in the principles of ecological restoration. *Environmental Monitoring and Assessment*, *190*, 113. https://doi.org/10.1007/s10661-018-6494-9

Park, H., Turner, N., & Higgs, E. (2018). Exploring the potential of food forestry to assist in ecological restoration in North America and beyond. *Restoration Ecology*, *26*, 284–293. https://doi.org/10.1111/rec.12576

Pascual, U., Balvanera, P., Díaz, S., Pataki, G., Roth, E., Stenseke, M., Watson, R. T., BaşakDessane, E., Islar, M., Kelemen, E., Maris, V., Quaas, M., Subramanian, S. M., Wittmer, H., Adlan, A., Ahn, S., Al-Hafedh, Y. S., Amankwah, E., Asah, S. T., … , Yagi, N. (2017). Valuing nature's contributions to people: The IPBES approach. *Current Opinion in Environmental Sustainability*, *26–27*, 7–16. https://doi.org/10.1016/j.cosust.2016.12.006

Peytchev, A., Peytcheva, E., Conzelmann, J. G., Wilson, A., & Wine, J. (2020). Modular survey design: Experimental manipulation of survey length and monetary incentive structure. *Journal of Survey Statistics and Methodology*, *8*, 370–384. https://doi.org/10.1093/jssam/smz006

Plieninger, T., Dijks, S., Oteros-Rozas, E., & Bieling, C. (2013). Assessing, mapping, and quantifying cultural ecosystem services at community level. *Land Use Policy*, *33*, 118–129. https://doi.org/10.1016/j.landusepol.2012.12.013

Russell Bernard, H. (2011). *Research methods in anthropology: Qualitative and quantitative approaches* (5th ed.). AltaMira Press.

Scholte, S. S. K., van Teeffelen, A. J. A., & Verburg, P. H. (2015). Integrating socio-cultural perspectives into ecosystem service valuation: A review of concepts and methods. *Ecological Economics*, *114*, 67–78. https://doi.org/10.1016/j.ecolecon.2015.03.007

Stephenson, J. (2008). The cultural values model: An integrated approach to values in landscapes. *Landscape and Urban Planning*, *84*, 127–139. https://doi.org/10.1016/j.landurbplan.2007.07.003

Stern, M. J., Dillman, D. A., & Smyth, J. D. (2007). Visual design, order effects, and respondent characteristics in a self-administered survey. *Survey Research Methods*, *1*, 121–138.

Tress, B., & Tress, G. (2001). Capitalising on multiplicity: A transdisciplinary systems approach to landscape research. *Landscape and Urban Planning*, *57*, 143–157.

Vannieuwenhuyze, J., Loosveldt, G., & Molenberghs, G. (2010). A method for evaluating mode effects in mixed-mode surveys. *Public Opinion Quarterly*, *74*, 1027–1045.

Wartman, P., Van Acker, R., & Martin, R. C. (2018). Temperate agroforestry: How forest garden systems combined with people-based ethics can transform culture. *Sustainability*, *10*, 2246. https://doi.org/10.3390/su10072246

Young, A. (Ed.). (1997). Agroforestry, soil management and sustainability. In *Agroforestry for soil management* (2nd ed., pp. 1–22). CAB International.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.