sustainability

MDPI

Article

A Flexible Inventory of Survey Items for Environmental Concepts Generated via Special Attention to Content Validity and Item Response Theory

John A. Vucetich [1,*], Jeremy T. Bruskotter [2], Benjamin Ghasemi [3], Claire E. Rapp [4], Michael Paul Nelson [4] and Kristina M. Slagle [2]

[1] College of Forest Resources and Environmental Science, Michigan Technological University, Houghton, MI 49931, USA
[2] School of Environment and Natural Resources, The Ohio State University, Columbus, OH 43210, USA; bruskotter.9@osu.edu (J.T.B.); slagle.44@osu.edu (K.M.S.)
[3] Human Dimensions of Natural Resources Department, Colorado State University, Fort Collins, CO 80523, USA; benjamin.ghasemi@colostate.edu
[4] Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA; claire.rapp@oregonstate.edu (C.E.R.); mpnelson@oregonstate.edu (M.P.N.)
* Correspondence: javuceti@mtu.edu

Abstract: We demonstrate how many important measures of belief about the environmental suffer from poor content validity and inadequate conceptual breadth (dimensionality). We used scholarship in environmental science and philosophy to propose a list of 13 environmental concepts that can be held as beliefs. After precisely articulating the concepts, we developed 85 trial survey items that emphasized content validity for each concept. The concepts' breadth and the items' content validity were aided by scrutiny from 17 knowledgeable critics. We administered the trial items to 449 residents of the United States and used item response theory to reduce the 85 trial items to smaller sets of items for use when survey brevity is required. The reduced sets offered good predictive ability for two environmental attitudes (R^2 = 0.42 and 0.46) and indices of pro-environmental behavior (PEB, R^2 = 0.23) and behavioral intention (R^2 = 0.25). The predictive results were highly interpretable, owing to their robust content validity. For example, PEB was predicted by the degree to which one believes nature to be sacred, but not by the degree of one's non-anthropocentrism. Concepts with the greatest overall predictive ability were *Sacredness* and *Hope*. Belief in non-anthropocentrism had little predictive ability for all four response variables—a claim that previously could not have been made given the widespread poverty of content validity for items representing non-anthropocentrism in existing instruments. The approach described here is especially amenable to incremental improvement, as other researchers propose more informative survey items and potentially important concepts of environmental beliefs we overlooked.

Keywords: environmental attitudes; content validity; item response theory; anthropocentrism; pro-environmental behaviors

check for updates

Citation: Vucetich, J.A.; Bruskotter, J.T.; Ghasemi, B.; Rapp, C.E.; Nelson, M.P.; Slagle, K.M. A Flexible Inventory of Survey Items for Environmental Concepts Generated via Special Attention to Content Validity and Item Response Theory. *Sustainability* **2024**, *16*, 1916. https://doi.org/10.3390/su16051916

Academic Editors: Fanli Jia and Genovaitė Liobikienė

Received: 15 October 2023
Revised: 6 February 2024
Accepted: 14 February 2024
Published: 26 February 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

A useful definition of sustainability is meeting human needs in a socially just manner without depriving ecosystems of their health [1]. Making substantial progress toward sustainability can depend on a critical understanding of that definition. Critical understandings include knowing, for example, what people believe about value-laden concepts such as non-anthropocentrism, and what people believe it means for an ecosystem to be healthy [1]. In this broad regard, there is a perennial need for robust survey instruments that are capable of quantifying beliefs about the environment. It is commonly understood that the robustness of an instrument depends on its empirical properties, as revealed by

factor analysis and measures of reliability [2]. No less important are a priori conceptual considerations that precede empirical evaluations of instrument performance, such as the development of survey items that adequately represent a precisely articulated underlying concept [3]. This consideration is sometimes associated with the term "content validity" [3] (but see [4]). Concerns about the content validity of instruments of environmental belief are widespread and include the following:

- Being inadequately rooted to theory (e.g., [5,6]).
- Having gaps between their underlying structures and presupposed measurement models (e.g., [7]; see also [8]).
- Being too difficult for many subjects to interpret (e.g., [9,10]).
- Having become outdated as environmental crises have deepened over the past four decades [5,11–13].
- Poor representation of the articulated underlying construct [14,15].

The broadest expression of this concern is that many instruments are plausibly not measuring what they purport to measure.

This study has two aims. The first is to detail the aforementioned concerns. The second aim is to prioritize content validity in developing an inventory of survey items of environmental beliefs. From the large set of trial items that we developed, we identified smaller sets of items using item response theory (IRT)—as opposed to the more commonly applied principles of classical test theory [16]. Finally, we assess the ability of these items to predict environmental behaviors, behavioral intentions, and two overarching attitudes about the environment.

2. A Priori Conceptual Considerations

The structure of Section 2 is as follows. Section 2.1 is a review of prior research [17], which indicates how most survey items used to represent (non-)anthropocentrism fail to do so because the wordings of the items do not reflect accepted conceptualizations of what (non-)anthropocentrism is. In other words, these items lack content validity. Sections 2.2 and 2.3 use the revised New Environmental Paradigm (NEP) [12] to illustrate how the statistically supported dimensionality of a survey instrument can be poorly aligned with existing concepts in environmental discourse, such as (non-)anthropocentrism or nature's fragility, and result in dimensions without any clear interpretation. After detailing those concerns for NEP, Section 2.4 demonstrates how those concerns are likely a common feature of many survey instruments pertaining to environmental attitudes. Section 2.1 through Section 2.4 give ample reason to consider the development of a survey instrument that focuses on content validity from the perspective of scholarly discourse in environmental science and philosophy. Finally, Section 2.5 explains how item response theory (IRT) pairs nicely with an interest in focusing on content validity.

2.1. Content Validity of Items

The importance of a priori conceptual considerations that should precede empirical evaluations of survey instruments is exemplified here by a qualitative method for assessing content validity. We do so by considering one dimension of NEP, i.e., "anti-anthropocentrism," which is represented by three survey items:

- Humans have the right to modify the natural environment to suit their needs. (2)
- Plants and animals have as much right as humans to exist. (7)
- Humans were meant to rule over the rest of nature. (12)

The numbers following each item within rounded parentheses are numbers assigned to each item by Dunlap et al. [12]. Anti-anthropocentrism is represented by disagreement with statements 2 and 12 and agreement with statement 7.

Because anthropocentrism and non-anthropocentrism are formal jargon from the academic field of environmental ethics, it is useful to assess the content validity of the aforementioned items in relationship to that scholarship, which is reviewed in [17]. According

to that scholarship, anthropocentrism is defined as the belief that *only* humans (and all humans) possess intrinsic value. What NEP refers to as anti-anthropocentrism is more commonly referred to as non-anthropocentrism, which is the belief that all humans *and* at least some portion of the nonhuman world possess intrinsic value. Intrinsic value is value that transcends use value or instrumental value (which is roughly equivalent to utility in the parlance of economics). If the *only* value of an object is its utility to another, then it does not possess intrinsic value. Humans have an obligation to treat possessors of intrinsic value with respect, fairness, or concern for their wellbeing or interests. A non-anthropocentrist can decide against the interests of one with intrinsic value (human or otherwise), but they cannot do so without good reason. To be non-anthropocentric does not determine how one will adjudicate instances of conflict among possessors of intrinsic value. Rather, to be non-anthropocentric only means that adjudication is required, as opposed to simply dismissing the conflict.

This account of non-anthropocentrism indicates that items in NEP, intended to represent non-anthropocentrism, are poor indicators of the concept articulated by environmental philosophers. For example, an anthropocentrist and a non-anthropocentrist can both agree with item 2 (above) because anthropocentrism is not about whether humans have a right to do anything, rather it is about whether there is a need to adjudicate a conflict between the wellbeing of humans and nonhumans.

An anthropocentrist and a non-anthropocentrist can also both agree with item 12, because non-anthropocentrism is not about whether humans were meant to rule nature. Non-anthropocentrism could allow for humans (a moral agent) to "rule" others, so long as the ruling is benevolent to the humans and nonhumans who possess intrinsic value. This describes our relationship with, for example, children, incapacitated humans, and domestic animals—all of whom possess intrinsic value. While we tend not to use the word "rule" to describe those relationships, the point is that "ruling" does not preclude the acknowledgment of intrinsic value.

Item 7 may be an acceptable representation of non-anthropocentrism, though non-anthropocentrism is not really—in a formal sense—about the "right to exist". Nevertheless, confirmatory factor analysis indicates that item 7 should be excluded from NEP (see below).

Concerns about the content validity of items intending to represent non-anthropocentrism are not isolated to NEP. An earlier review identified 42 items, representing the vast majority of items used to assess anthropocentrism or non-anthropocentrism. Of those items, 39 had significant conceptual shortcomings [17]. For example, in some published studies, agreement with these statements is intended to represent anthropocentrism:

- Nature is important because of what it can contribute to the pleasure and welfare of humans.
- The worst thing about the loss of the rain forest is that it will restrict the development of new medicines.
- The best thing about camping is that it is a cheap vacation.

And agreement with these statements has been taken to indicate non-anthropocentrism:

- Forests give us a sense of peace and wellbeing.
- Forests rejuvenate the human spirit.
- Forests let us feel close to nature.
- I need time in nature to be happy.

Neither agreeing nor disagreeing with those statements is indicative of anthropocentrism or non-anthropocentrism. The shortcomings rise from easy-to-overlook aspects of intrinsic value. For example, a person can believe that an entity possesses intrinsic value, but not have a positive affective response to that entity, and vice versa. Similarly, deriving personal benefit from an entity is no indication of whether one believes that entity possesses intrinsic value. These and other commonly misapprehended features of intrinsic value are detailed in [17].

These critiques would be of modest import if it were not possible to develop and identify better alternatives. But better alternatives seem to exist, such as:

- Nature is important *only* for what it can provide to humans.
- Wild animals are *only* valuable if people get to utilize them in some way.
- The *only* real value of an ecosystem (forest, lake, or river) is for the products and services they provide to humans.

In those statements, the word "only" is especially important for ensuring that the items align with the definition of anthropocentrism.

Statements representing non-anthropocentrism are more difficult to develop, because that concept is literally defined by what it is not. Consequently, there are many ways to be non-anthropocentric. Nevertheless, a simple statement that better represents non-anthropocentrism is:

- The needs of nature are important, even when meeting those needs is of no benefit humans.

Other statements that can better distinguish anthropocentrism and non-anthropocentrism are:

- The only reason to conserve nature is to meet human interests.
- Nature should be treated fairly, with concern for its own needs.

What seems to make these items adequate is their vague reference to "nature".

2.2. Dimensionality

In addition to content validity, the hypothesized dimensionality of environmental beliefs is also an important a priori conceptual consideration. To illustrate, consider the five hypothesized dimensions of the NEP:

- Reality of limits to growth (O).
- Fragility of nature's balance (O).
- Anti-anthropocentrism (O).
- Rejection of human exemptionalism (R).
- Possibility of an ecocrisis (R).

Dimensions marked with (O) were part of the original NEP [18], and dimensions marked with (R) were added for the revised NEP [12]. Principle component analysis indicates that much of the variance in data collected through the revised NEP is explained by four, unlabeled dimensions that are not well aligned with the hypothesized dimensions [12]. Confirmatory factor analysis (CFA) also indicates that data collected through the revised NEP are a poor fit to the hypothesized five-dimensional structure [2]. Post hoc assessment of that CFA suggests three dimensions that were labeled:

- Limits to growth.
- Concern about ecological damage.
- Anti-anthropocentrism.

Here, "anti-anthropocentrism" is represented by items 2, 4, 8, 10, and 12 (Table 1). One concern is that items 4, 8, and 10 have nothing to do with anthropocentrism or non-anthropocentrism. Furthermore, CFA leads to the conclusion that item 7 should be excluded from the revised NEP (Table 1), even though it was originally hypothesized to represent anti-anthropocentrism and is more closely aligned with the concept of non-anthropocentrism than other items.

Table 1. Items in the revised New Environmental Paradigm, numbered as in Dunlap et al. [12].

Hypothesized Dimensions	Empirically Supported Dimensions	Survey Items
Limits	Limits	We are approaching the limit of the number of people the Earth can support. (1)
Limits	Limits	The Earth is like a spaceship with very limited room and resources. (11)

Table 1. *Cont.*

Hypothesized Dimensions	Empirically Supported Dimensions	Survey Items
Anti-ant	Anti-ant	Humans have the right to modify the natural environment to suit their needs. (2)
Anti-ant	Anti-ant	Humans were meant to rule over the rest of nature. (12)
Anti-Exempt	Anti-ant	Human ingenuity will ensure that we do not make the Earth unlivable. (4)
Balance	Anti-ant	The balance of nature is strong enough to cope with the impacts of modern industrial nations. (8)
Eco-crisis	Anti-ant	The so-called "ecological crisis" facing humankind has been greatly exaggerated. (10)
Anti-Exempt	Concern	Despite our special abilities, humans are still subject to the laws of nature. (9)
Balance	Concern	When humans interfere with nature it often produces disastrous consequences. (3)
Eco-crisis	Concern	Humans are seriously abusing the environment. (5)
Eco-crisis	Concern	If things continue on their present course, we will soon experience a major ecological catastrophe. (15)
Limits		The Earth has plenty of natural resources if we just learn how to develop them. (6)
Anti-Exempt		Humans will eventually learn enough about how nature works to be able to control it. (14)
Anti-ant		Plants and animals have as much right as humans to exist. (7)
Balance		The balance of nature is very delicate and easily upset. (13)

The first column indicates the hypothesized dimension to which each item belongs, and the second column indicates the dimension to which each item was assigned following confirmatory factor analysis [2]. Note the dimension labels in the first and second columns: Limits—limits to growth; Anti-ant—anti-anthropocentrism; Anti-Exempt—anti-exemptionalism; Balance—fragility of nature's balance; Eco-crisis—possibility of an eco-crisis; Concern—concern about ecological damage.

2.3. Null Hypotheses of Dimensionality

The epistemological value of a CFA—as with any science leaning on the falsification of null hypotheses—depends on the details of the null hypothesis. The rejection of a strong null hypothesis contributes more knowledge than the rejection of a weak null hypothesis [19]. This is why the hypothesized dimensionality of any construct deserves careful a priori consideration. When a goodness-of-fit test for a CFA indicates that data are a poor fit to a hypothesized dimensional structure, then it matters whether the data are a poor fit to a conceptually robust hypothesized dimensionality or a poor fit to an ad hoc and difficult-to-interpret dimensionality. In other words, when CFA rejects a hypothesized dimension that is not well articulated, one does not really know what is being rejected. Thus, it matters whether the rejected dimensionality was well developed and the extent to which it is supported by scholarly discourse on environmental thought. The concern may be expressed as a rhetorical question: Does the rejection of a hypothesized dimension indicate a problem with the underlying concept that supports the dimension or with the survey items used to assess that dimension? Statistical inference does not answer that question.

Consider those ideas in the context of NEP. All that was documented about the grounding of the hypothesized dimensions of NEP is reprinted in the Appendix A. That documentation is sufficiently sparse to warrant concluding that due attention had not been given to the breadth or nuance of concepts in environment discourse or the content validity of items intending to represent such concepts. The concern is that CFA (conducted by [2]) may have rejected a hypothesis that was not especially well grounded in relevant theory, and therefore of lesser value to the growth of knowledge pertaining to how humans mentally organize various beliefs about the environment.

These concerns are not mitigated by using CFA to infer an alternative dimensionality (after rejecting an a priori null hypothesis of dimensionality) or by using EFA to infer the di-

mensions of environmental beliefs in an ad hoc manner. In other words, reverse engineering an underlying construct from the factor analysis of a small collection of statements would seem too often to be fraught with conceptual uncertainty. For example, consider the NEP items belonging to the empirically supported dimension labelled anti-anthropocentrism (Table 1, 2nd column). That dimension may measure something of import, but it is not clear what that underlying construct would be.

2.4. Hypothesized Dimensionality of Other Instruments

Cruz and Manata [2] conducted a systematic search of the literature and identified 18 significant instruments for measuring environmental attitudes (typically from cultures that are western, educated, industrialized, rich and democratic, i.e., WEIRD). Of the 18 instruments, they identified five as most important, including NEP. Next, we review the dimensionality of the remaining four instruments as a basis for suggesting—as others have—that concerns about content validity are widespread. The critiques that follow do not discount genuine insights rising from these instruments. They are only intended to inspire further development.

First, Maloney et al. [20] developed an instrument with three dimensions: self-reported behaviors, behavioral intentions, and environmental attitudes. CFA indicates the appropriateness of considering environmental attitudes as a single dimension [2]. This is not surprising given that all of the items are focused on a particularly narrow aspect of human–nature relationships, i.e., affective response to pollution:

- It genuinely infuriates me to think that the government doesn't do more to help control pollution of the environment.
- I become incensed when I think about the harm being done to plant and animal life by pollution.
- I'm usually not bothered by so-called "noise pollution".
- When I think of the ways industries are polluting, I get frustrated and angry.
- The whole pollution issue has never upset me too much since I feel it's somewhat overrated.

Note, Maloney et al. [20] originally included 10 items. These five items result in a better fit to the CFA and higher reliability [2]. In any case, adequately representing *general* beliefs about the environment would seem to need items covering a broader swath of ideas than affective response to pollution.

Second, Weigel and Weigel [21] developed a survey instrument by conceiving survey items "focus[ed] on a wide range of conservation and pollution issues". No other theoretical grounding is offered. While Weigel and Weigel [21] supposed the items to represent a single dimension, CFA indicates two dimensions [2]. One dimension is labeled "rejection of industrial status quo" and is represented by items such as:

- Although there is continual contamination of our lakes, streams, and air, nature's purifying processes soon return them to normal.
- Predators such as hawks, crows, skunks, and coyotes which prey on farmers' grain crops and poultry should be eliminated.

The second is labeled "concern about pollution" and is represented by items such as:

- The federal government will have to introduce harsh measures to halt pollution since few people will regulate themselves.
- I'd be willing to make personal sacrifices for the sake of slowing down pollution even though the immediate results may not seem significant.

The concerns here are similar to concerns with Maloney et al. [20]. One might expect an instrument representing general beliefs about the environment to include a greater breadth of ideas and ideas that are less specific than these two dimensions.

This instrument also seems compromised by concerns about reverse engineering a construct from a few statements. For example, it is not at all clear how the two items for the dimension "rejection of industrial status quo" relate to that construct. Similarly, though more subtly, if one wanted to design survey items about pollution, one might avoid items

that invoke extraneous and potentially distracting ideas (such governmental regulation and personal sacrifice). These subtle shortcomings risk measurement error and likely contribute to the notoriously limited capacity to predict more precise attitudes and behaviors from general beliefs.

The third study from Cruz and Manata [2] to review is Lounsbury and Tornatzky [22], who developed a survey instrument:

> "us[ing] a panel of four judges composed of two psychology graduate students and two members of a state environmental action organization, a 78-item attitude questionnaire was constructed. The items represented a variety of content domains including issues of overpopulation, pollution, economic materialism, conservation, and environmental action."

Lounsbury and Tornatzky [22] used statistical analyses (described only as cluster analysis) to reduce the number of items to 12. CFA indicates the presence of three dimensions, which led to reducing the instrument to 10 items [2]. The result is a three-dimensional scale consisting of these items.

Dimension 1: Concern for environmental degradation.

- * The news media have exaggerated the ecological problem.
- If mankind is going to survive at all, environmental pollution must be stopped.
- I am worried about future children's chance of living in a clean environment.
- * We shouldn't worry about environmental problems because science and technology will solve them before very long.

Dimension 2: Concern for environmental action.

- People should buy (and return) beverages only in returnable containers.
- People should use less detergent than the manufacturer recommends to help preserve water quality.
- * There is nothing wrong with using electric can openers, electric pencil sharpeners, and electric toothbrushes.
- * Putting a brick in one's toilet to conserve water is a dumb idea.

Dimension 3: Concern for overpopulation.

- Every couple in America should try not to have more than two children.
- Overpopulation is a major source of environmental problems today.

Items marked with an asterisk are reverse-coded. Prima facie, this scale seems reasonable—though items in dimension 2 refer to dated technologies and behaviors in a way that likely contributes to measurement error. Nevertheless, one wonders whether a more robust scale would have resulted from a more systematic consideration of the concepts in environmental discourse.

The third study from Cruz and Manata [2] to review is Schultz [23], who developed a survey instrument whose underlying constructs pertain to motivations that might lead a person to be concerned for the environment. The hypothesized dimensionality of this instrument was supported by Cruz and Manata [2] and slightly adjusted (by removing two items) to improve reliability. The result is an instrument prompted by this statement: "I am concerned about environmental problems because of the consequences for _____," where the blank stands for one of the following objects, organized into one of three dimensions:

Dimension 1: Biospheric concern.

Plants;
Marine life;
Birds;
Animals.

Dimension 2: Egoistic concern.

Me;
My health;

My future.

Dimension 3: Social–altruistic concern

All people;
Children.

Each object is rated on a scale of 1 (not important) to 7 (supreme importance). This scale is particularly well suited to the values–belief–norms model [24,25] and norm-activation theory [26]. Another strength of this scale is its precise focus on one's motivation to behave in certain ways. Yet, that precision also means this scale is likely too narrow to adequately represent different kinds of beliefs about the environment.

To summarize the preceding review, poor content validity seems to be a widespread feature of survey research on environmental beliefs. That limitation does not negate the genuine insights rising from such research. But it does represent an important opportunity to develop survey items by giving due attention to (1) content validity and (2) a list of concepts that better represent the breadth of important ideas in environmental discourse. In the next section, we explain why we attend those ideas in conjunction with item response theory (IRT).

2.5. Item Response Theory

Before explaining the rationale for using item response theory (IRT), we provide a brief summary of IRT, which can begin by observing that the response to any survey item rises from both traits of a respondent and the traits of the item. IRT aims to account for both aspects of a response by modeling a sample of responses to a set of survey items with two kinds of statistic [27–29]. The first kind of statistics are threshold parameters, which indicate where on the scale of the latent trait an item provides information. Items with low thresholds provide the most information about those exhibiting low levels of the latent trait, and items with high thresholds provide the most information about those exhibiting high levels of the latent trait. Threshold parameters are useful for scale development because they indicate whether a set of items provides information from across an appropriate range of the scale.

For multilevel responses, such as the Likert-scale items used in the present study, IRT can rely on the graded response model, which is a variant of the two-parameter logistic model that treats each item response category as a binary response [30]. If the item has a five-point response set, then there are four threshold parameters, which indicate the value on the scale of the latent trait for which a respondent would select a response of "2" or greater 50% of the time, a "3" or greater 50% of the time, a "4" or greater 50% of the time, or a "5" 50% of the time. There is no threshold parameter for 1 or greater, because respondents always select a 1 or greater.

The second statistic common to IRT is a discrimination parameter (DP), which indicates the extent to which a response to an item provides precise information about where on the scale a respondent is with respect to the latent trait. DPs also indicate how closely associated an item is with the latent trait. As such, DPs are loosely analogous to factor loadings in a factor analysis.

We used IRT for several reasons. First, threshold parameters are valuable for judging the steep trade-off between two important properties of a survey: a brief survey and a survey that includes enough items to cover an appropriate range of the measured scales. Carefully tending that trade-off is important for environmental beliefs, because those beliefs include many concepts, which would seem to require many scales (see Section 3.1).

Second, IRT pairs nicely with our primary interest, to develop an inventory of concepts and develop survey items for those concepts while giving due attention to content validity. To see how, observe that survey items and the data they produce may be evaluated for various psychometric properties, such as factorial structure and various aspects of validity and reliability. Typically, it is not possible to design one survey instrument that maximizes all properties at the same time [31]. More specifically, in the domain of environmental beliefs, a set of survey items can have a high level of content validity or discriminant validity,

but having both at the same time seems elusive (as indicated in Section 2). Furthermore, many survey instruments have been developed by prioritizing, for example, an optimal factorial structure and acceptable values of Cronbach's alpha. Because the statistical tools of IRT do not emphasize those aspects of psychometry, it is a convenient framework from which to focus on content validity. More generally, most existing scales of environmental beliefs were designed from principles of classical test theory. As such, there is value in better understanding how IRT can be used in survey research on environmental beliefs.

3. Formulation of Dimensions and Trial Items

3.1. Formulation of Concepts

To develop an inventory of survey items for general beliefs about the environment, we began by developing a proposed list of concepts based on our collective interdisciplinary understanding of environmental scholarship. We drew on each of our native disciplines—environmental ethics, environmental psychology, and environmental science—as well as our decades-long experience with interdisciplinarity amongst those fields.

To develop this list of concepts, we alternated between independent thinking and group thinking (among the co-authors). More specifically, we held 60–90 min meetings on a weekly basis over a period of more than three months, where each meeting was preceded by independent thinking on some prescribed aspect of environmental beliefs, such as brainstorming a list of metaphysically oriented concepts or judging whether two concepts should be merged into one. We met to share and discuss the results of our independent thinking.

After developing a list of concepts—along with a succinct, precise articulation of each concept—we shared that list with 11 scholars, each with significant training and experience in some aspect of environmental scholarship, such as environmental science or environmental philosophy. We asked each scholar to reflect on whether we missed any important concept and whether the concepts were adequately articulated. We discussed the reflections of these scholars and revised the list accordingly. To the extent that labels can be useful, this method is comparable to a modified Delphi approach [32,33] (see also [15]).

Throughout this process, we judged each concept according to the following:

- Whether it represents ideas about human–nature relationships that are more specific than basic values (e.g., Schwartz Value Inventory) and transcend attitudes about specific environmental issues (e.g., attitudes about carbon taxes).
- Whether it represents a reasonably important topic in environmental discourse, as indicated perhaps by its importance in the media or scholarly writings about the environment.
- Whether it is sufficiently distinct from other proposed concepts, in the sense that beliefs about one concept do not impose a significant rational constraint on beliefs about another concept (e.g., anthropocentrism and non-anthropocentrism would not be separate concepts because those ideas are entirely defined in terms of each other).

We did not presuppose that any particular concept would antecede or predict more specific environmental attitudes or pro-environmental behaviors or behavioral intentions. We considered such ideas as hypotheses to be tested, rather than to be presupposed.

3.2. Description of Concepts

The concepts that we developed are detailed in the following paragraphs, and each concept is indicated by italics.

Non-anthropocentrism is a belief and fundamental theme in environmental ethics that was defined and detailed in Section 2.1.

Comfort is the degree to which thoughts of nature trigger the affective responses of feeling comforted or threated. That stimulus–response relationship has a deep evolutionary history [34]. Affective (or emotional) responses to a stimulus differ from reasoned, belief-based concepts. One important distinction is that emotions are processed in different parts

of the brain and occur prior to reasoning in decision making [35]. *Comfort* is also widely appreciated as a basic concept in environmental psychology [36,37].

Connectedness is the extent to which one believes humans and nature are essentially separate or one-and-the-same. That is a perennial topic in environmental philosophy [38] and can be an important basis for evaluating environmental policies (e.g., [39]). The topic is often treated by environmental philosophers as metaphysical, in the sense of not being readily resolved by objective, empirical evaluation in much in the same way that one cannot empirically distinguish half-empty and half-full glasses of water.

Dependency is the extent to which one believes human wellbeing is resilient to environmental degradation. This concept is rooted in ideas from environmental economics, i.e., strong and weak sustainability, which pertain to the extent that one assumes that humans can replace natural capital, goods and services upon being depleted [40]. Strong and weak sustainability have been an important basis for evaluating a wide range of environmental policies since the 1980s.

Stability is the extent to which one believes nature, in its healthy state, tends to be stable or dynamic. *Fragility* is the extent to which one believes nature is fragile (as opposed to resilient) to human impacts. The relevance of both concepts is indicated by the perennial literature pertaining to the "balance of nature" [41–45].

Hope is the extent to which one believes there is hope for a good outcome in the relationship between humans and nature. Hope is a widely studied, general construct in psychology [46], but it also has important particular manifestations in the context of environmental thought (e.g., [47–49]).

Doubting Others is the extent that one believes the fate of human–nature relationships depend on the actions of others, as opposed to one's own actions. This belief is important from various scholarly perspectives. From a psychological perspective, it is associated with response efficacy, which can antecede pro-environmental behavior [50]. While we could have labelled this concept "response efficacy," we refrained from doing so because *Doubting Others* is a specific kind of response efficacy.

From the perspective of environmental sociology, *Doubting Others* is associated with understanding the extent to which one views an environmental issue as a tragedy of the commons that requires collective action [51]. From the perspective of ethics, this belief is associated with the degree to which one acts on the basis of consequentialism, as opposed to virtue-based or deontologically based frameworks of ethical decision making (Morrell and Dahlmann [52] for a review of those ethical frameworks). We considered whether *Hope* and *Doubting Others* were sufficiently related to merge into a single concept. Because a case can be made for either decision, we decided more could be learned by distinguishing these ideas at this stage of development.

Sacredness is the degree that one believes nature is sacred in the sense of having a character beyond what is material or secular. This concept receives perennial attention from various perspectives, including theologies of organized religions and beliefs associated with Indigenous culture and Neopaganism (e.g., [53–57]).

Holism is short for ecological holism. While the term is associated with a variety of subtly distinguishable beliefs [58], we use the term to refer to the belief that ecological collectives (such as a species or biological community) are living individuals. In other words, holism is a belief that objects, such as a forest, are a kind of individual organism. The concept is closely associated with the metaphysical views of some prominent ecologists [59–63] and Deep Ecology [64].

Animism is also associated with a variety of subtly distinguishable beliefs [65]. We use the term to represent the degree to which one attributes life to nonliving things [66], more precisely, to entities treated by Western science as nonliving, such as rocks, water, and air. The connection between beliefs about animism and human–nature relationships is widely appreciated in the scholarly literature in anthropology [67] and religion [68].

Nature's breadth is the extent to which one believes that various objects, such as golf courses and logged forests, are natural even though they may have been impacted or

constructed by humans. The importance of this concept is associated with perennial concern over the logical fallacy known as appeal-to-nature [69]. A heuristic example of appeal-to-nature reasoning is as follows: a logged forest is unnatural; therefore, it is not good. Alternatively, an unlogged forest is natural; therefore, it is good. Appeals-to-nature are important because many human minds are drawn to such reasoning, and the reasoning is considered poor for two reasons: (1) there is no logical connection between naturalness and goodness; (2) in too many cases, the boundary between natural and unnatural is arbitrary. While that boundary may be arbitrary, it is perceived by many. This concept aims to assess how broad (inclusive) one takes the category of natural to be.

Shared traits is best understood by first attending to the notion of anthropomorphism, which has been defined variously by psychologists. One representative definition is as follows [70]:

> "anthropomorphism is the human tendency to assign human characteristics, motives, behaviors and abilities to nonhuman entities, particularly animals. Anthropomorphized wildlife characters are pervasive in human life and discourse—in stories and myths, material representations, advertisements, analogies, and normative depictions of culture."

Another representative definition is as follows [66]:

> "Psychologists have used the term anthropomorphism rather loosely to describe everything from mistaken inferences about nonhuman agents to almost any kind of dispositional inference about a nonhuman agent, definitions that do not fit with the actual dictionary definition of attributing "human characteristics or behavior to a god, animal, or object".

Those accounts of anthropomorphism are useful, but fail to account for important nuance, as illustrated by these instances of anthropomorphism:

- Cows having four-chambered hearts;
- Cows enjoy reading Proust;
- Cows can experience joy;
- Cows have souls.

While each acknowledgement is an instance of anthropomorphism (according to commonly applied definitions), the acknowledgements also differ importantly from each other. In particular:

- The first instance is an objective, empirical claim that is uncontroversially true.
- The second instance is an objective, empirical claim that is uncontroversially false.
- The third instance is an objective, empirical claim widely supported by experts in the psychology of affect in nonhuman animals, but not so widely accepted by non-experts.
- The fourth claim is a metaphysical claim whose truth-value is not evaluated by empirical analysis.

With that context, *Shared Traits* is the extent to which one attributes nonhuman animals with traits that have at times and by some been considered distinctively human traits, but scientific inquiry suggests they are at least plausibly attributable to at least some nonhumans. *Shared Traits* may also be related to one's empathic capacity, insomuch as perceiving similarity in another is an important predictor of empathy and sometimes an antecedent for moral consideration [71].

3.3. Formulation of Trial Items

We developed trial survey items for each concept using an approach similar to that used for developing the list of concepts. In total, we developed approximately 110 items. We asked seven graduate students from various fields of environmental research to critique the trial items on the following bases: (1) simplicity of grammar and vocabulary and unambiguousness; (2) close adherence to the underlying concept. Based on that evaluation, we settled on 85 trial items that are presented in Tables 2–14.

Table 2. *Sacredness* items and their item response theory statistics.

Item Code	DP	Threshold Values				Item Wording
		1	2	3	4	
Sa1	1.4	−3.02	−2.27	−1.31	0.09	Nature is sacred.
Sa2	0.31	−6.83	−3.69	−1.12	1.71	Nature is no more than a bunch of creatures each trying to survive in the world.
Sa3	1.96	−2.55	−1.91	−0.57	0.5	To look at nature is to look at the divine.
* Sa4	1.49	−2.19	−1.37	0.09	1.45	Meditation is an essential way to understand nature.
* Sa5	3.05	−2.49	−1.86	−0.93	0.23	I feel blessed when I am in nature.
Sa6	3.96	−2.08	−1.47	−0.43	0.55	I have an important connection with nature that I can feel in my soul.
* Sa7	3.02	−1.74	−1.19	−0.26	0.83	My relationship with nature is importantly spiritual.
Sa8	0.36	−6.23	−1.75	2.51	6.44	The connections in nature are complicated, but not divine (mystical).
Sa9	0.67	−1.99	−0.77	1.97	3.03	Nature is better understood through, science: meditation

Sacredness is the extent to which one believes nature is sacred, in the sense of having a character beyond what is material or secular. The response set for each item was a 5-point scale from "strongly agree" to "strongly disagree." Except Sa9, which was structured as a semantic differential, where "science" and "meditation" were at the two poles of the differential. DP is the discrimination parameter. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *.

Table 3. *Hope* items and their item response theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
H1	0.73	−3.36	−0.86	1.25	3.7	I am optimistic that humans will soon begin to live sustainably.
* H2	2.19	−1.95	−1.00	−0.12	1.23	Humans appear doomed to ruin the planet.
* H3	1.5	−1.23	0.04	1.11	2.29	It is too late to fix our most dire environmental crises.
* H4	2.15	−1.36	−0.48	0.38	1.54	Humans are too weak-willed to solve environmental crises.
H5	1.8	−1.93	−1.09	0.06	1.26	Solving environmental crises depends on government and business, but they are unable or unwilling to do so.
H6	0.25	−11.5	−5.65	−0.11	4.18	Human wellbeing is unlikely to be adversely impacted by climate change, because science and engineering will find ways for us to adapt.

Hope is the extent to which one believes there is hope for a good outcome in the relationship between humans and nature. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2.

Table 4. *Doubting Others* items and their item response theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
* DO1	1.56	−0.36	0.64	1.69	2.55	How much faith do you have in government officials trying to protect the environment?
* DO2	2.58	−1.18	−0.41	0.33	1.36	If government officials tried to protect the environment, how much difference would it make?
DO3	1.47	−0.66	0.49	1.73	2.72	How much faith do you have in business leaders trying to protect the environment?
* DO4	2.97	−1.47	−0.53	0.21	1.16	If business leaders tried to protect the environment, how much difference would it make?
DO5	1.03	−2.22	−0.53	1.18	2.81	How much faith do you have in regular people to do the right things to protect the environment?
* DO6	1.54	−2.37	−1.12	0.12	1.05	If regular people tried to protect the environment, how much difference would it make?

Doubting Others is the extent that one believes that the fate of human–nature relationships depends on the actions of others, as opposed to one's own actions. The response set for each item was as follows: very little, a little, a moderate amount, a lot, a great deal. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2.

Table 5. *Nature's Breadth* items and their item response theory statistics.

Item	DP	Threshold Values			Item Wording
		1	2	3	
* NB1	3.18	−1.69	−0.72	0.35	Golf course
* NB2	1.78	−1.69	−0.06	1.67	Neighborhood park
* NB3	1.19	−1.64	−0.43	0.66	Logged forest
NB4	−0.33	−2.08	−6.81	−10.29	Unlogged forest
NB5	0.9	−3.49	−1.41	0.49	Genetically modified plant
NB6	0.42	−1.02	4.48	10.06	Garden plant
NB7	−0.5	−2.95	−7.2	−12.35	Wild plant

Nature's Breadth is the extent to which one believes various objects, such as golf courses and logged forests, are natural even though they may have been impacted or constructed by humans. These items were presented as a matrix, led by the following question: "How would you describe each item listed below?" The rows of the matrix consisted of the phrases listed below and the columns were the response set, whose wording was: perfectly natural, mostly natural, slightly natural, not natural at all. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2.

Table 6. *Connectedness* items and their item response theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
Cn1	0.56	−4.00	−0.57	1.31	3.89	Humans depend on nature, but we are still fundamentally separate from nature.
* Cn2	1.53	−2.15	−0.99	0.02	1.44	Humans and nature are one-in-the-same.
* Cn3	4.63	−2.14	−1.43	−0.63	0.55	Humans and nature are intimately linked, like family.
Cn4	0.91	−4.16	−1.3	−0.48	0.85	To harm nature is to harm ourselves.
* Cn5	1.04	−4.98	−2.06	0.1	2.16	How connected to nature do you feel?

Connectedness is the extent to which one believes humans and nature essentially separate or one-and-the-same. The response set for Cn5 was as follows: not at all, a little connected, moderately connected, deeply connected, one with nature. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2.

Table 7. *Comfort* items and their item response theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
* Cm1	4.36	−2.22	−1.72	−1.21	−0.35	comforted: threatened
* Cm2	3.7	−2.32	−1.79	−1.28	−0.36	uneasy: peaceful
* Cm3	4.55	−2.23	−1.6	−1.09	−0.31	content: miserable
Cm4	2.49	−2.5	−2.03	−1.14	−0.22	nurturing: dangerous
Cm5	1.97	−2.62	−2.22	−1.27	−0.28	in need of taming and domination: life-giving and healing

Comfort is the degree to which thoughts of nature trigger the affective responses of feeling comforted or threated. These items were semantic differentials, where the prompting statement is as follows: "When I am in nature, I feel:" The two poles of the differential are indicated by the words and phrases below. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2.

Table 8. *Shared Traits* items and their Item Response Theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
ST1	0.8	−3	−2.03	0.04	2.42	Tool use is widespread among mammals and birds.
ST2	1.75	−2.5	−1.99	−1.15	0.36	There are animals (beside humans) that use language.
* ST3	1.63	−3.2	−2.55	−0.9	0.36	Fish can experience pain.
ST4	2.69	−2.58	−2.07	−0.99	0.13	Simple emotions (such as joy and fear) are widespread among animal species.

Table 8. *Cont*.

Item	DP	Threshold Values 1	2	3	4	Item Wording
ST5	2.65	−2.43	−1.79	−0.65	0.6	Complicated emotions (like embarrassment and pride) can be found among at least some animals with complicated social lives, like wolves and baboons.
* ST6	1.89	−3.07	−2.38	−0.67	0.81	Squirrels differ among themselves in terms of personality.
* ST7	1.23	−2.82	−1.78	0.26	1.88	Flies experience at least a limited kind of consciousness.
ST8	2.07	−2.93	−2.32	−0.82	0.68	There are many examples of mammal species with minds that give those animals intention (i.e., a choice to behavior one way or another).
ST9	1.27	−3.65	−2.61	0.07	1.48	Deer can remember what happened yesterday.

Shared Traits is the extent to which one attributes nonhuman animals with traits that have at times and by some considered distinctively human traits, but scientific inquiry suggests are at least plausibly attributable to at least some nonhumans. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are the same as those in Table 2. For future surveys, we recommend replacing ST6 with slightly simpler wording: "Squirrels often differ from each other in terms of personality".

Table 9. *Fragility* items and their item response theory statistics.

Item	DP	Threshold Values 1	2	3	4	Item Wording
* F1	1.85	−2.69	−1.61	−0.85	0.73	The balance of nature tends to be fragile and easily upset by humans.
F2	2.19	−2.03	−1.18	−0.52	0.9	Damage to nature tends to heal quickly.
F3	2.13	−2.72	−1.45	−0.69	0.65	Nature tends to be delicate and easily harmed by humans.
* F4	1.84	−2.23	−0.84	−0.01	1.38	Nature tends to be resilient to human impacts.
F5	1.88	−2.62	−1.33	−0.5	0.9	Harms to Nature's health tend not to be easily healed.
F6	2.29	−2.55	−1.57	−0.77	0.69	Natural environments tend to be easily degraded by human use.
F7	2.07	−2.55	−1.5	−0.72	0.83	Losses to nature's well-being tend not to be easily restored.
* F8	2.54	−1.91	−1.1	−0.39	0.75	Nature tends to recover quickly from harms caused by humans.
F9	2.54	−2.23	−1.75	−0.87	0.05	Nature's health is:—easy for humans to ruin:difficult for humans to ruin
F10	3.47	−1.75	−1.45	−0.67	0.12	When humans harm nature, the damage tends to be:—easy to repair:difficult to repair
F11	2.66	−1.73	−1.33	−0.61	0.22	Nature tends to recover from human harms:—quickly:slowly
F12	1.37	−2.31	−1.21	0.21	1.16	Nature tends to recover from human harms:—on its own:only with strict protections
F13	1.79	−2.21	−1.48	−0.37	0.68	The balance of nature tends to be:—delicate:sturdy
F14	1.9	−2.23	−1.44	−0.24	0.82	The balance of nature tends to be:—fragile:robust

Fragility is the extent to which one believes nature is fragile (or resilient) to human impacts. The last 6 items are semantic differentials. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2.

Table 10. *Dependency* items and their item response theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
D1	0.83	−3.54	−2.39	−0.93	0.99	Human well-being is threatened by environmental degradation.
D2	2.51	−2.09	−1.2	−0.62	0.2	Overfishing is not a problem because humans are able to find other sources of food.
D3	3.52	−1.94	−1.39	−0.83	−0.01	Overlogging is not a concern because forests don't provide anything that cannot be substituted with technology.
D4	1.77	−3.27	−2.28	−1.32	−0.2	Nature requires special protection because it is an important and irreplaceable source of human well-being.
D5	1.34	−3.26	−2.19	−1.05	0.38	We are depleting nature's bounty in ways that are essential for human well-being and not easily replaced.
* D6	1.6	−2.18	−1.13	−0.2	1.02	Technological innovation will maintain human well-being in the face of environmental degradation.
* D7	2.12	−2.12	−0.99	−0.14	0.73	Recent environmental changes are not generally all that bad for human well-being.
D8	1.89	−1.77	−0.86	−0.25	0.49	The adverse impact of climate change and pollution tends to be exaggerated.
* D9	2.98	−1.94	−1.35	−0.82	0.08	Water pollution is not much of a concern because water treatment plants provide clean water.

Dependency is the extent to which one does not believe human wellbeing is resilient to environmental degradation. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2.

Table 11. *Non-Anthropocentrism* items and their item response theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
Ant_a1	0.89	−4.2	−3.04	−1.30	0.56	Nature has inherent value, beyond any benefit that humans get from using nature.
Ant_a2	1.43	−3.2	−2.41	−1.36	0.18	Nature should be treated with at least some concern for its well-being, irrespective of whether humans benefit as a result.
Ant_a3	2.55	−1.43	−0.92	−0.34	0.36	Nature is valuable only when humans benefit from using it.
Ant_b1	12.50	−1.23	−0.89	−0.63	−0.11	Rivers, lakes and oceans exist only to serve human needs.
Ant_b2	15.77	−1.24	−0.92	−0.62	−0.12	Forests and prairies exist only to serve human needs.
Ant_b3	9.98	−1.31	−0.93	−0.65	−0.17	Wildlife exists only to serve human needs.

Non-Anthropocentrism is the belief that only humans possess intrinsic value. Other details are as in Table 2. The notation for these items reflects our decision to retain all six items but consider them to be separate dimensions (*Non-Anthropocentrism* a and b). See text for details. Ant_a2 might be improved for future surveys by striking the phrase "at least some".

Table 12. *Animism* items and their item response theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
* Ani1	2.02	−2.14	−1.34	−0.74	0	Is a rock that you can hold in your hand alive?
* Ani2	39.32	−1.06	−0.49	−0.16	0.51	Is the water that you could pour into a glass alive?
* Ani3	3.96	−1.06	−0.36	0.09	0.7	Is the air that you breathe alive?

Animism is the degree to which one attributes life to entities treated by Western science as nonliving, such as rocks, water, and air. The response set for these items was: definitely yes, probably yes, not sure, probably no, definitely no. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2.

Table 13. *Stability* items and their item response theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
* Stb1	1.01	−1.25	0.05	1.17	2.15	dynamic: constant
* Stb2	1.27	−0.69	0.89	2.26	3.3	never-changing: ever-changing
* Stb3	27.96	−0.57	0.31	1.19	1.55	always fluctuating: forever fixed

Stability is the extent to which one believes that nature, in its healthy state, tends to be stable or dynamic. These items were semantic differentials, where the prompting statement is, "Nature in its healthy state tends to be:" The two poles of the differential are indicated by the words and phrases below. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2.

Table 14. *Holism* items and their item response theory statistics.

Item	DP	Threshold Values				Item Wording
		1	2	3	4	
* Hol1	0.28	−2.16	0.25	3.92	6.93	a collection of living things (trees, birds, insects, etc.): a unified whole, like a single organism
* Hol2	25.02	−1.74	−1	−0.27	0.62	Oceans are composed of many individual living things (individual fish of all kinds). But is an ocean itself and as a whole an individual living thing?
* Hol3	2.85	−1.78	−1.18	−0.4	0.66	Earth is inhabited by many individual living things. But is Earth itself and as a whole an individual living thing?

Holism is the belief that ecological collectives, such as a species or biological community, are living individuals. Hol1 was structured as a semantic differential, where the prompting statement is, "A forest is:" The two poles of the differential are indicated by the two phrases above that are separated by a colon. For the remaining items, the response sets were as follows: definitely yes, probably yes, not sure, probably no, definitely no. The reduced set of items whose predictive ability was assessed (as described in Section 5.3) are marked with *. Other details are as in Table 2. For future surveys, we recommend restructuring Hol1 to match the other items, using the statement, "Forests are composed of many individual living things (plants, insects, birds, mammals). But is a forest itself an individual living thing".

4. Sampling and Survey

In July and August of 2022, we used the Qualtrics platform to administer an online survey to a panel of 449 adults (>18 years) residing in the United States that approximates the distribution of age, sex, and income at the time of the 2020 census (Appendix B). Approximately 50% of the sample comprised people living in rural communities, which is high compared to census data which indicate that about 20% of U.S residents live in rural communities. That difference does not compromise any of the conclusions that we draw below.

5. Analysis and Results

5.1. Item Reduction via Item Response Theory

To assess the items and concepts, we used the statistical program IRTPRO v4.2 and a graded response model [30] using the Bock–Aitkin estimation method. The principles of classical test theory suppose that the multiple, redundant items are "the root of precision" (p. 57 in [72]). In item response theory, however, the basis for judging the adequacy for a set of items is the discrimination parameters and threshold values. We use that basis with an aim of reducing the number of items for each concept to three or four.

More specifically, to guide the item reduction process, we used discrimination parameters (DPs) and threshold values. Higher DPs are better than lower DPs, ceteris paribus. While there is no hard rule, DPs <0.64 can be concerning and scores >1.34 are more than adequate [73]. With respect to threshold values, sets of survey items whose threshold values cover the range [–2.0, 2.0] are good when the interest is to distinguish people across a broad range of the latent traits [29]. When DP- and threshold-based criteria provided an equivocal basis for item reduction, we then considered other properties, such as the grammatical simplicity of an item or a post hoc evaluation of content validity.

The results below are ordered so that similarly performing concepts appear, more or less, in succession. These results are also detailed in Tables 2–14. In those tables, each

trial item is identified with notation such H1 for item 1 from *Hope* and Sa2 for item 2 from *Sacredness.* We use such notation beginning in the next sentence.

Sacredness: The three items with the highest DPs were Sa5 (3.1), Sa6 (4.0), and Sa7 (3.0) (Table 2). In the previous sentence, and from this point forward, unlabeled numbers in rounded parentheses are DPs. The range of threshold values for that set of items was [–2.5, 0.8]. However, post hoc consideration of Sa6′s content validity suggests that it might overlap with *Connectedness*. Furthermore, removing Sa6 does not reduce the range of thresholds. For those reasons, we removed it from consideration.

The next most useful item was Sa4, whose DP was 1.5. Its inclusion with Sa5 and Sa7 increases the range of threshold values to [–2.5, 1.5]. These items (Sa4, Sa5, Sa7) are likely to be a useful scale for *Sacredness*, but there would be value in future research that developed items with good discrimination at the top end of the scale.

Hope: The three items with the highest DPs were H2 (DP = 2.2), H4 (DP = 2.2), and H5 (DP = 1.8) (Table 3). Item H3 also had an acceptably high DP (1.5). Because H3 and H5 had similar DPs, we recommend selecting H3 instead of H5 for two reasons. Namely, H5 has more complex grammar and its content validity might be obscured by the prominent reference to government and business. The range of threshold values for that set of items (H2, H3, H4) was [–2.0, 2.3]. Other items had unacceptably low DPs or did not lead to an increased range of threshold values. While future research might develop items with higher DPs at the higher and lower ends of the scale, this set of items is at least a good starting place for assessing *Hope*.

Doubting Others: The items DO2 (3.0) and DO4 (2.6) had significantly higher discrimination than the remaining items (Table 4). Three other items (DO1, DO3, and DO6) had acceptable and similar DPs. Of those three, one had threshold values that covered the lower end of the scale (DO6), and another had threshold values that covered the higher end of the scale (DO1). For these reasons, we recommend retaining four items to represent *Doubting Others.* This set of four items have threshold values covering the range [–2.4, 2.6].

Nature's Breadth: We eliminated several items for having DPs less than one (NB4, NB5, NB6, and NB7, Table 5). The remaining three items had acceptable DPs, i.e., NB1 (3.2), NB2 (1.8), and NB3 (1.2). The range of threshold values for those items was [–1.7, 1.7]. There would be value in developing items with good discrimination at the lower and upper portions of the scale in future research.

Connectedness: We eliminated two items for having low DPs. They are Cn1 (0.6) and Cn4 (0.9) (Table 6). The remaining three items had acceptable DPs (1.0, 1.5, and 4.6). The range of threshold values for this set of three items was [–5, 2.2].

Comfort: All five items performed similarly in the sense of having acceptable DPs and having similar ranges of threshold values (Table 7). As such, we picked the three items with the highest DPs, Cm1 (4.4), Cm2 (3.7), and Cm3 (4.6). As a set, these items had a range of [–2.3, –0.4]. There would be value in developing items with good discrimination at the lower and upper portions of the scale in future research.

For each of the next four concepts, the trial items exhibited a trade-off in the sense that DPs were negatively correlated to both the range of threshold scores and maximum threshold score. In other words, items with better discrimination tended to represent the lower portion of the scale (i.e., individuals with low values of trait), but not the upper portion.

For these concepts, the first step we took in item reduction was to eliminate items with low discrimination. Then, we identified an item with high discrimination (and tending to cover the low end of the scale) and two items that had better coverage at the high end of the scale (and tending to have lower discrimination). When that second criterion identified more than two items, we selected the two items with simpler grammar or better content validity.

While each of the next three scales are likely to be useful, there would be value in developing items with higher discrimination at the top end of the scale in future research.

Shared traits: For these nine trial items (Table 8), the DPs were negatively correlated to both the range of threshold scores ($r = 0.87 \pm 0.10$ SE) and the score of the upper threshold ($r = 0.82 \pm 0.14$). We eliminated ST1 from consideration for having lower discrimination (0.8). The item that we selected with higher discrimination was ST6 (1.9). The items that we picked with coverage at the higher end of the scale were ST3 (1.6) and ST7 (1.2). The range of threshold values for this set of three items is [–3.2, 1.9]. Note that we recommend selecting ST3 over ST8 because ST8 has considerably more complex grammar. For example, the Flesch–Kincaid index for sentence complexity indicates that ST3 is readable down to 4th grade level, but ST8 is rated at a 12th grade or higher.

Fragility: For these 14 trial items (Table 9), the DPs were negatively correlated to both the range of threshold scores ($r = 0.84 \pm 0.09$) and the score of the upper threshold ($r = 0.78 \pm 0.12$). None of the items needed to be eliminated for having unacceptably low discrimination. The item that we selected with higher discrimination was F8 (2.5). The items that we picked with coverage at the higher end of the scale were F1 (1.9) and F4 (1.8). The range of threshold values for this set of three items is [–2.7, 1.4].

Dependency: For these nine trial items (Table 10), the DPs were negatively correlated with both the range of threshold scores ($r = 0.90 \pm 0.08$) and the score of the upper threshold ($r = 0.62 \pm 0.25$). We eliminated D1 from consideration for having low discrimination (0.83). The item that we selected with higher discrimination was D6 (3.0). The items that we picked with coverage at the higher end of the scale were D6 (1.6) and D7 (2.1). The range of threshold values for this set of three items is [–2.2, 1.0].

Non-anthropocentrism: Three items had a wider range of threshold values, but notably lower discriminations (Table 11). In particular, the items Ant_a1 (0.9), Ant_a2 (1.4), and Ant_a3 (2.6) had threshold values whose range was [–4.2, 0.56]. These three items were also characterized by the word "nature" being the subject of the statements.

The other three items had more specific subjects (e.g., rivers or forests). Those items had especially high DPs (>9), but an especially narrow range of threshold values, i.e., [–1.3, –0.2].

Because we could not say in advance whether items with a subject that is generic (nature) or specific (rivers or forests) would have better predictive ability, we treated the first three items as being indicative of a dimension that we labeled *Non-anthropocentrism a* and the second three items as indicative of *Non-Anthropocentrism b.*

For each of the next three dimensions, we developed only three trial items, because we were unable to think of additional items that would have high content validity without being especially redundant with other items. Consequently, there is no need to reduce the number of items. But there is value in evaluating their performance.

Animism: All three items had DPs > 2 (Table 12). The range of threshold value for this set items covered the range [–2.1, 0.7]. There would be value developing items with good discrimination at the upper portions of the scale in future research.

Stability: All three items had DPs > 1 (Table 13). Item Stb3 had especially high discrimination (28) and a range of threshold values that covered the highest end of the scale, i.e., [–0.6, 1.6]. The range of threshold values for this set of three items was [–2.2, 3.3]. There would be value in developing items with better discrimination in future research. We guess that the higher discrimination of Stb3 is associated with word choices that are more meaningful to a general audience.

Holism: Item Hol1 was eliminated because its DP was 0.3 (Table 14). The remaining two items had high discrimination, i.e., Hol2 (25) and Hol3 (2.9). The range of thresholds for this pair of items was [–1.8, 0.7].

The item with poor discrimination was a semantic differential for which many survey participants may have seen as a false dichotomy. We recommend restructuring that item to be like the items that performed better, so that it reads: *Forests are composed of many individual living things (individual trees and animals). But is a forest itself an individual living thing.*

In any case, there would be value in developing items with good discrimination at the lower and upper portions of the scale in future research.

5.2. Second-Order Factor Analysis

In the context of IRT, second-order factor analysis is often importantly moot. However, there is value in performing such analysis given our attention to content validity and the interpretation of underlying concepts. In particular, second-order exploratory factor analysis (EFA) can suggest whether or how the many dimensions of a broad construct (environmental beliefs) can be hierarchically arranged into a fewer (statistically parsimonious) number of interpretable factors. We consider the results of EFA for this purpose.

For emphasis, this insight cannot be gained from confirmatory factor analysis (CFA), which assesses the degree to which the data deviate from some preconceived hypothesis about the data structure. We are not claiming that it would be wrong to apply CFA to the data. Rather, we are claiming that EFA better fulfills the intended purpose. Namely, knowing if the list of concepts (Section 3.2) is organized in the minds of non-experts (general sample of survey respondents) in any particular way that is interpretable. In other words, CFA essentially provides only a yes or no answer to a question about whether data fit the one particular structure that is specified in the null hypothesis. But EFA goes beyond answering that yes or no question by also suggesting how the data might be best structured.

More specifically, for each survey participant, we calculated an average response for the three or four items representing each concept to represent a score for each concept. We then performed EFA with R and the commands fa(fm ="ml", rotate = "oblimin") and fa.parallel() on the reduced set of items. A scree plot suggests that three dimensionality is appropriate (Figure A2, Appendix C). The three dimensions, the variance explained, and the loadings are presented in Table 15. There was no significant cross loading, except readers should note that *Nature's Breadth* loads 0.42 on Factor A and −0.40 on Factor B. The strength of the loading is similar for the two factors. We were unable to develop compelling ad hoc descriptions for any of the factors, given the dimensions that loaded highest for each factor (see also Section 6.3).

Table 15. Factor loading for indices of various dimensions of environmental belief based on survey items identified by item response theory as being most informative.

	Factor		
Dimension	**A**	**B**	**C**
Dependency	(0.82)	0.01	−0.02
Non-anthropocentrism b	(0.78)	−0.06	−0.10
Non-Anthropocentrism a	(0.67)	0.22	−0.04
Fragility	(0.63)	0.02	0.28
Nature's breadth	(0.42)	−0.40	0.16
Connectedness	0.08	(0.78)	−0.08
Sacredness	−0.03	(0.65)	0.18
Doubting others	0.03	(−0.53)	−0.07
Comfort	0.27	(0.42)	−0.02
Hope	0.01	0.01	(−0.48)
Animism	−0.19	0.23	(0.39)
Shared traits	0.20	0.37	0.24
Holism	0.04	0.28	0.35
Stability	−0.20	0.04	0.04

The factors explained 47% (A), 37% (B), and 16% (C) of the variance. Loadings in parentheses have the highest loadings for each factor. Dimensions highlighted in dark gray had the most predictive ability (according to Tables 16–19). Dimensions with light gray highlighting had the next most predictive ability.

Table 16. Items representing pro-environmental behaviors and behavioral intentions.

Category	Survey Items
Behaviors	. . .how much money have you donated to environmental organizations (please enter an estimated dollar amount)? . . .how frequently did you volunteer your time for an environmental organization? [never, once or twice, three or four times, once a month, once every couple weeks, most weeks] . . .how often did you participate in grass roots environmentalism, such as attend a rally or protest? [never, once or twice, three or four times, once a month, more than once a month] . . .how often did you express your views about the environment to a politician or government official (e.g., by writing a letter)? [never, once or twice, three or four times, once a month, more than once a month]
Intentions	. . .significantly limit nonessential shopping and consumerism? . . .significantly reduce your consumption of electricity and fuel? . . .limit your consumption of red meat? . . .significantly increase the portion of your diet that comes from local sources? . . .significantly increase how often you encourage others to act in favor of the environment?

Items representing behaviors were preceded by the phrase, "In the past year". Text within square brackets are response sets. Items representing intentions were preceded by the phrase, "In the upcoming year, how likely are you to. . ." The response set for those items was a 7-pt scale from "extremely likely" to "extremely unlikely".

Table 17. Regression models indicating the predictive ability of various dimensions of environmental belief (n = 424).

k	R^2	Sacredness	Doubting Others	Connected	Non-Anthro a	Hope	Depend	Holism	Nonant B	Comft
1	0.30	1.57 ± 0.12 *								
2	0.38	1.24 ± 0.12 *	−0.79 ± 0.11 *							
3	0.41	0.96 ± 0.13 *	−0.62 ± 0.11 *	0.74 ± 0.14 *						
4	0.45	0.92 ± 0.12 *	−0.61 ± 0.11 *	0.64 ± 0.14 *	0.54 ± 0.12 *					
5	0.45	0.90 ± 0.13 *	−0.61 ± 0.11 *	0.64 ± 0.14 *	0.53 ± 0.12 *	−0.17 ± 0.10				
6	0.45	0.91 ± 0.13 *	−0.63 ± 0.11 *	0.62 ± 0.14 *	0.44 ± 0.14 #	−0.17 ± 0.10	0.14 ± 0.11			
7	0.45	0.89 ± 0.13 *	−0.61 ± 0.11 *	0.62 ± 0.14 *	0.43 ± 0.15 #	−0.16 ± 0.10		0.12 ± 0.10	0.12 ± 0.10	
8	0.46	0.86 ± 0.13 *	−0.62 ± 0.11 *	0.60 ± 0.14 *	0.43 ± 0.15 *	−0.17 ± 0.10		0.12 ± 0.10	0.11 ± 0.10	−0.21 ± 0.21

Each row is the result of best subsets regression models, where k is the number of predictor variables, and the columns are the predictors. Each cell reports a regression coefficient and standard error. An * refers to p-values < 10^{-3}, and # refers to $p < 0.01$. Lightly shaded cells indicate $p < 0.05$, and dark-shaded cells are $p \geq 0.05$. The response variable for these models is an 11-point scale (0–10) that is a response to the following question: "Compared to other interests in your life, how interested are you in issues pertaining to nature and the environment?".

Table 18. Regression models indicating the predictive ability of various dimensions of environmental belief (n = 442).

k	R^2	Fragility	Non-Anthro b	Hope	Depend	Nature's Breadth	Non-Anthro a	Doubting Others	Animism
1	0.26	−1.51 ± 0.12 *							
2	0.35	−1.07 ± 0.13 *	−0.80 ± 0.10 *						
3	0.38	−0.93 ± 0.13 *	−0.86 ± 0.10 *	0.48 ± 0.11 *					
4	0.40	−0.72 ± 0.14 *	−0.70 ± 0.11 *	0.50 ± 0.10 *	−0.45 ± 0.14 #				
5	0.41	−0.71 ± 0.14 *	−0.65 ± 0.11 *	0.50 ± 0.10 *	−0.42 ± 0.14 #	−0.32 ± 0.13			
6	0.41	−0.67 ± 0.14 *	−0.56 ± 0.12 *	0.48 ± 0.10 *	−0.36 ± 0.14	−0.34 ± 0.13	−0.32 ± 0.16		
7	0.42	−0.62 ± 0.14 *	−0.57 ± 0.12 *	0.48 ± 0.11 *	−0.38 ± 0.14 #	−0.39 ± 0.14 #	−0.29 ± 0.16	0.14 ± 0.11	
8	0.42	−0.63 ± 0.15 *	−0.57 ± 0.12 *	0.50 ± 0.10 *	−0.36 ± 0.14	−0.39 ± 0.14 #	−0.29 ± 0.16	0.17 ± 0.11	0.11 ± 0.08

The response variable for these models is an 11-point scale (0–10, "terrible" to "perfect, no need for improvement") that is a response to the question, "How would you rate humans' overall treatment of the environment?" An * refers to p-values < 10^{-3}, and # refers to $p < 0.01$. Other details are as in Table 16.

Table 19. Regression models indicating the predictive ability of various dimensions of environmental belief (n = 447).

k	R^2	Doubting Others	Sacred	Nature's Breadth	Fragility	Hope	Connected	Holism	Depend
1	0.13	−0.54 ± 0.07 *							
2	0.18	−0.41 ± 0.07 *	0.39 ± 0.08 *						
3	0.20	−0.35 ± 0.07 *	0.37 ± 0.08 *	−0.29 ± 0.08 *					
4	0.21	−0.39 ± 0.07 *	0.40 ± 0.08 *	−0.23 ± 0.09 #	−0.19 ± 0.07				
5	0.22	−0.39 ± 0.07 *	0.38 ± 0.08 *	−0.22 ± 0.08 #	−0.22 ± 0.07 #	−0.14 ± 0.06			
6	0.23	−0.36 ± 0.07 *	0.31 ± 0.08 *	−0.21 ± 0.09	−0.22 ± 0.07 #	−0.14 ± 0.06	0.17 ± 0.09		
7	0.23	−0.35 ± 0.07 *	0.30 ± 0.08 *	−0.22 ± 0.09	−0.22 ± 0.07 #	−0.13 ± 0.06	0.14 ± 0.09	0.07 ± 0.07	
8	0.23	−0.35 ± 0.07 *	0.29 ± 0.08 *	−0.20 ± 0.09	−0.18 ± 0.09	−0.12 ± 0.07	0.17 ± 0.10	0.07 ± 0.07	−0.06 ± 0.08

The response variable for these models is an index of pro-environmental behaviors. An * refers to p-values < 10^{-3}, and # refers to $p < 0.01$. Other details are as in Table 15.

5.3. Predictive Ability

A survey instrument is often said to have predictive validity if it is correlated with another well-established instrument representing the same or similar underlying construct [74]. Assessing this kind of validity can be important if there are doubts about what an instrument is measuring. Predictive validity is not the goal of the analysis described in this section. (We write about predictive validity in Section 6.4.)

Rather, the purpose here is two-fold. The first purpose is to assess hypotheses about the degree to which the environmental concepts that we operationalized are predictive of measures for environmental behaviors, behavioral intentions, and two overarching attitudes about the environment. This predictive ability is taken as evidence for or against those hypotheses, not as evidence for the quality of the items' ability to measure what they purport to measure. The second purpose is to assess predictive ability while being mindful of the often-severe constraints on survey length, which will sometimes prevent an analyst from presenting as many survey items as we have developed. We fulfill these purposes with best subsets regression, described just below. For clarity, our purpose here is not to test any particular behavioral theory (e.g., value-belief-norm theory), though we discuss the relevance of such theories in Section 6.2.

To perform this analysis, we used regression and four dependent variables to quantify the predictive ability of the survey items that we developed. The dependent variables were:

- Responses to a survey item about the overall importance of environmental issues (Table 16 for details).
- Responses to a survey item about how well humans treat nature (Table 17 for details).
- A scale based on responses to five items pertaining to pro-environmental behavioral intentions (PEBIs).
- A scale based on responses to four items pertaining to pro-environmental behaviors (PEBs).

We developed items about PEBs and PEBIs to represent behaviors that seemed likely to impact the environment (sensu, [75]). The items also represented behaviors that we believe were not entrained by habit for most participants. The PEBIs also represented sociopolitical actions (e.g., writing a politician about an environmental issue) that can be accurately recalled and reported. The items representing PEBs and PEBIs are presented in Table 16. Other details pertaining to those items are presented in Appendix D.

We used best subsets regression—i.e., the function regsubset() in program R—which finds the best model with k predictors. We examined models for k = [1, 2, . . .8]. The model results are reported in Tables 17–20.

Drawing attention to the models with only statistically significant predictors, the most predictive models explained 44% and 40% of the variance for the survey items pertaining to the importance and treatment of the environment (Tables 17 and 18). For PEBIs and PEBs, the most predictive models comprised of only statistically significant predictors explained 19% (PEBIs) and 24% (PEBs) of the variance (Tables 19 and 20).

Table 20. Regression models indicating the predictive ability of various dimensions of environmental belief (n = 447).

k	R^2	Sacred	Hope	Connected	Animism	Depend	Doubting Others	Fragility	Shared Traits
1	0.17	0.55 ± 0.06 *							
2	0.20	0.53 ± 0.06 *	−0.18 ± 0.05 *						
3	0.22	0.41 ± 0.07 *	−0.18 ± 0.05 *	0.25 ± 0.07 *					
4	0.23	0.37 ± 0.07 *	−0.16 ± 0.05 #	0.24 ± 0.07 *	0.11 ± 0.04				
5	0.24	0.41 ± 0.06 *	−0.15 ± 0.05 #		0.11 ± 0.04 #	0.15 ± 0.05 #	−0.16 ± 0.05 #		
6	0.25	0.35 ± 0.07 *	−0.15 ± 0.05 #	0.16 ± 0.07	0.11 ± 0.04 #	0.13 ± 0.05 #	−0.13 ± 0.06		
7	0.25	0.36 ± 0.07 *	−0.17 ± 0.05 #	0.15 ± 0.07	0.11 ± 0.04 #	0.18 ± 0.06 #	−0.14 ± 0.06	−0.09 ± 0.07	
8	0.25	0.35 ± 0.07 *	−0.16 ± 0.05 #	0.13 ± 0.08	0.11 ± 0.04	0.17 ± 0.06 #	−0.14 ± 0.06	−0.10 ± 0.07	0.09 ± 0.08

The response variable for these models is an index of self-reported pro-environmental behavioral intentions. An * refers to p-values < 10^{-3}, and # refers to $p < 0.01$. Other details are as in Table 15.

The regression results suggest that a plausible ranking of the concepts' predictive ability would be as follows.

Most predictive:

- *Sacredness* was a strong predictor for three of the four responses (import, PEBI, PEB).
- *Hope* was a strong predictor for two responses (treatment, PEBI), an important predictor for a third (PEB), and possibly a weak predictor for a fourth response (import).
- *Doubting Others* was a strong predictor for three responses (import, PEBI, PEB) and possibly a weak predictor for a fourth response (treatment).
- *Dependency* was a strong predictor for two responses (treatment, PEBI), and a weak predictor for two responses (treatment, PEB).

Moderately predictive:

- *Fragility* was a strong predictor for one response (treatment), an important predictor of another response (PEB), and a weak predictor for another response (PEBI).
- *Connectedness* was a strong predictor of two responses (import, PEBI).
- *Nature's breadth* was a strong predictor of one response (PEBI) and an important predictor of another (treatment).
- *Non-anthropocentrism a* was a strong predictor of one response (import) and an important predictor of another response (treatment).

Minimally predictive:

- *Non-anthropocentrism b* was an important predictor of one response (treatment) and perhaps a weak predictor of another response (import).
- *Animism* was perhaps a weak predictor of two responses (treatment, PEBI).

Least predictive:

- *Comfort* was perhaps a weak predictor of one response (import).
- *Holism* was perhaps a weak predictor of one response (PEB).
- *Shared traits* was perhaps a weak predictor of one response (PEBI).
- *Stability* was not a predictor of any response.

This ranking is intended to be no more than a qualitative summary of Tables 17–20; it is not a claim about the ability of survey items to measure what they purport to measure, nor is it a broad claim about the general importance of those aspects of environmental beliefs to predict other phenomena (see Section 6.4). Rather, they are claims about the dataset that we analyzed. For additional context, see Table A1, Appendix E for a matrix of bivariate correlations among the variables.

6. Discussion

Much research on environmental beliefs—including their antecedents and consequences—is based on survey instruments whose development rose from giving considerable attention to the empirical properties of the survey instruments. The empirical properties receiving the most attention include the correlations among items within hypothesized dimensions (Cronbach's

alpha) and correlations among sets of items that represent hypothesized dimensions (factor analysis). Attention to empiricism is well warranted, and the insights produced by such survey instruments are genuinely valuable. Yet, the strong focus on those empirical properties also represents an opportunity to explore insights rising from the development of survey items that prioritize content validity.

6.1. Interpretation

A distinctive benefit of survey items developed via prioritizing content validity is the increased interpretability of results. For example, we found that people who more strongly believe that nature is sacred (*Sacredness*) and are more hopeful about nature's future (*Hope*) are more likely to exhibit stronger PEBI (Table 19). Furthermore, the extent to which one expresses PEBI was not related to either the degree to which one expresses non-anthropocentrism (*Non-anthropocentrism*) or the degree to which one holds a holistic view of nature (*Holism*). Those conclusions are importantly interpretable in the sense of indicating which kind of environmental beliefs are and are not predictive of PEBI.

That level of interpretability contrasts, for example, with prior research indicating a positive association between one's score on the NEP instrument and PEBI or PEB (see Table 3 in Cordano et al. [76]). That result is not so highly interpretable because the construct that underlies the NEP is not well defined, and the dimensions identified by factor analysis are not readily interpretable (Section 2). Similarly, prior research has concluded that non-anthropocentrism is weakly related to PEB or PEBI (e.g., [77]). However, such conclusions are not especially warranted insomuch as they rely on instruments of non-anthropocentrism with poor content validity.

The interpretive power of the items we developed applies not only to PEBs and PEBIs, but also to overarching attitudes about the environment. For example, the results presented here translate to a relatively crisp interpretation of why some people are more interested in environmental issues than others (Table 17). In particular, the results indicate that people who see nature as sacred (*Sacredness*) and see humanity as deeply connected to nature (*Connectedness*) express more interest in environmental issues.

Being hopeful was a strong predictor of PEBIs, and the direction of that relationship is noteworthy: being more hopeful was associated with weaker PEBIs (Table 20). The relationship between hope and PEBs was similar, though not as strong (Table 20). Furthermore, people who were more hopeful tended to think that environmental issues were less important among sociopolitical issues (Table 17). Finally, being more hopeful about the environment was strongly predictive of thinking that humans treat the environment well (Table 18). The strength of those relationships supports the emerging notion that environmental hope and its compliment—fatalism—are important among various kinds of environmental belief (e.g., [78]). The direction of the relationships involving hope indicate that its effect many be complicated and nuanced [79]. These conclusions about hope are further supported by noting the important predictive ability of *Doubting Others*—which is likely akin to environmental fatalism. *Doubting Others* had a strong inhibitory effect on PEB and a moderate inhibitory effect on PEBI (Tables 19 and 20).

6.2. Relationship to Psychological Models of Behavior

Much survey research on pro-environmental behaviors is contextualized by several general models of behavior, such as the value–norms–belief theory (VBN; [80]), norm activation theory (NAT; [26]), and cognitive hierarchy theory (CHT; [81]). The inventory of survey items developed here can be related to such models. For example, the most obvious points of connection may include *Sacredness* being a value (sensu, VBN, and CHT) and *Doubting Others* being an awareness of consequences (sensu, NAT). Please note that we underscore the phrases that are formal jargon for VBN, NAT, or CHT on their first appearance.

Other points of connection may be more complicated. For example, *non-anthropocentrism* is well characterized as a value, but it may also be considered an ascription of responsibility

(sensu, NAT), at least in the sense that non-anthropocentrism determines whom one is responsible for treating fairly.

Fragility is what many environmental philosophers would consider a metaphysical belief, as opposed to an empirical belief or ethical belief (about how one ought to behave). As such, *Fragility* it is not well characterized as a value or a belief (as that term is used with VBN), though it may have important properties of a value, such as being established early in life and not readily changed. Furthermore, *Fragility* may also be an antecedent of awareness of consequences, but not ascription of responsibility. Other concepts, such as *Dependency* and *Connectedness,* are also likely to have complicated and indefinite mappings onto general models of behavior.

The relevance of these indefinite mappings rises from much research in environmental psychology focusing on the explanatory power of those psychological models. While valuable, that focus may come at the expense of giving less attention to developing insight from attending to content validity or concepts from the rich literature on environmental philosophy that do not map precisely onto existing psychological models of behavior.

Additional context is provided by the role that values play in value–norms–belief theory and cognitive hierarchy theory. Both theories predict that several processes intervene on the relationship between values and PEB (or PEBI), resulting in significant attenuation of the relationship between values and behaviors. In agreement with those theories, *Non-anthropocentrism* was a weak predictor of PEBI and PEB (Tables 19 and 20). However, *Sacredness* was a strong predictor of PEB and PEBI, and *Connectedness* was a strong predictor of PEBI. While the broad nature of *Connectedness* makes it easy to theorize that its relationship to PEBI ought to be attenuated by many intervening processes, the apparent predictive ability *Connectedness* suggests otherwise. Of course, additional research is required to confirm the predictive ability of these concepts.

6.3. Cognitive Hierarchies and Heuristics

Because statistical parsimony is greatly valued in survey research, second-order factor analysis can be valuable for assessing hierarchical relationships among the concepts of environmental beliefs. Interestingly, factor analysis did not produce an especially interpretable set of factors (Table 15). For example, seemingly unrelated concepts were placed in the same dimension (e.g., *Sacredness* and *Doubting Others*) and seemingly related concepts were separated into different dimensions (e.g., *Hope* and *Doubting Others*). Perhaps the hierarchy of concepts implied by factor analysis rises from some unrecognized organizing feature of the human mind. However, it is also plausible that these concepts are not hierarchically organized in the mind. That prospect should not be surprising, because the concepts we assessed are not generally recognized as hierarchically related in environmental discourse.

The potential lack of hierarchical organization and the strong association between PEBI and certain generic beliefs (Section 6.2) is consistent with the following suppositions:

- Survey participants respond to items while relying primarily on System I thinking that is, fast, intuitive, and heavily reliant on heuristics.
- Some environmental beliefs (such as *Sacredness* or *Connectedness*) function as heuristics and strongly influence certain kinds of PEBI. The PEBIs most influenced by such heuristics might be for behaviors not yet entrained by habit. That kind of PEBI happens to be the kind that we presented to participants (Table 16).

Regardless of the merit of that explanation, some explanation seems required for (1) the widespread tendency for factor analysis to fail to find highly interpretable hierarchical organizations of environmental beliefs, and (2) the prospect that some broad, generic beliefs can be strong predictors of certain kinds of PEBI. Neither circumstance would seem to be adequately explained by VBN, NAT, or CHT.

6.4. The Theory-Ladenness of Measurement

In Section 5.3, we assessed hypotheses about environmental beliefs' relationships to environmental behaviors, behavioral intentions, and two overarching attitudes about the

environment. That assessment gives occasion to review concerns about the relationship between predictive validity and the theory-ladenness of measurement [82,83].

Concerns about the theory-ladenness of a measurement can be assessed by comparing the measurements of two independently developed instruments designed to measure the same phenomena [84]—but only if the two instruments were designed to measure the same phenomenon. This is the kind of comparison that is relevant for assessing predictive validity of a new instrument.

However, quantifying the ability of one scale to predict another is not necessarily an indication of validity. For example, quantifying the ability of *Connectedness* to predict PEBI provides virtually no insight about whether the *Connectedness* is valid measure of what it purports to measure or of any other concept.

The validity of the items developed here depends strongly on qualitative assessments of content validity and the assumption that variations in the latent trait causally produce variations in the measurement (i.e., item responses). Anchoring validity to those conditions is consistent with frameworks for philosophy of science known as realism [85] and moderate forms of operationalism—which emphasizes that the meaning of a concept is intimately tied to the method by which the concept is measured [86].

Our point is the following: anchoring validity to the qualitative assessment of content validity represents an important manifestation of a measurement's theory-ladenness. But anchoring validity to other procedures is no less theory-laden. Furthermore, contemporary treatments of measurement's theory-ladenness no longer treat it as a threat to knowledge or aim to eliminate it [83]. The goal, rather, is to understand its effect on the meaning of scientific inferences. To that end, it is useful that much of the theory-ladenness of our measures is laid bare, in the sense that each concept is precisely defined and the relationship between those definitions and each survey item is transparent.

6.5. Flexible Inventory

The inventory of items developed here is not intended to be a fixed or final. The purpose of a survey should guide the selection of an appropriate set of items. For example, if one had a compelling hypothesis that involves, for example, *Holism* or *Animism,* then those items should be included in a survey, aside from their expected predictive ability.

The selection of items may also be affected by practical considerations. For example, if one wants to include the concepts that presently appear most predictive and can afford to include 10 items in a survey, then one would likely include the 10 items for *Sacredness, Hope*, and *Doubting Others.* If one can afford 16 items, then they might also include the items for *Dependency* and *Fragility.* If one can afford to include 40 items in a survey, then it may pay to include items from each of the 13 concepts. Doing so will lead to a more refined sense of which environmental beliefs are most predictive for samples representing different sociodemographic or cultural groups.

We also expect others will provide good reasons to add to the list of 13 concepts that we developed, and then assess the relationship between new and existing concepts. And we expect others to propose new items within each concept. Doing so may lead to a wider range of threshold values in cases where the range is currently less than ideal (Section 6.6).

6.6. Threshold Ranges

For about half of the concepts, the range of threshold values did not cover the top end of the scales. The relevance of this result may be understood by an analogy with testing students in an educational setting. The purpose of some tests is to distinguish students according to a skill, including the identification of students with the best skills. Other times the purpose is to determine what portion of students possess a certain trait, such as an ability to perform a skill at a specified level. In the former case, an appropriately wide range of threshold values is especially important. In the latter case, content validity is especially important.

Instruments of environmental beliefs may also serve different purposes. For example, one may want to estimate the portion of population believing that nature is sacred. In such cases, the items' content validity is especially important, and the range of threshold values will be of secondary importance. Consider that purpose and the first item for *Sacredness* (Sa1 in Table 2), which is the following statement: "Nature is Sacred". That item has the highest content validity and adequate discrimination (DP = 1.4), but its range of thresholds spanned only the lower end of the scale (–3.0 to 0.1). If the purpose is to estimate the frequency of that belief, then the low range of threshold values may not be concerning.

However, if the purpose of an instrument requires accurate and precise measurement at the top end of the scale, then items like Sa1 would be inadequate. For example, better coverage that the top end of a scale like *Sacredness* would add important predictive capacity for PEBs if people who see nature as especially sacred are the ones most engaged in PEB.

6.7. Discrimination/Threshold Trade-Off

The items for several concepts exhibit a trade-off between the discrimination parameter and the range of threshold parameters. The most severe trade-off was among items for *Non-anthropocentrism* (Table 11). Those six items fell into two groups, each with three items. One group had high DPs and a low range of threshold values; the other group had low DPs and a wide range of threshold values.

The trade-off may be related to the items' content validity. In particular, the items whose subject was "nature" had low DPs, and items with more specific subjects (rivers, forests, and wildlife) had much higher DPs. The vagueness of "nature" may lend itself to varied interpreted by respondents. For five or six of the six items, the subjects were ecological collectives (nature, forests, etc.), as opposed to individual, nonhuman organisms (e.g., an oak tree or a lion). These post hoc considerations suggest that these items have good content validity for a particular kind of non-anthropocentrism, i.e., the acknowledgement of intrinsic value in ecological collectives. Respondents who score high on our scale of non-anthropocentrism may also acknowledge the intrinsic value of individual, nonhuman organisms, but these items are unable to make that distinction.

6.8. Dimensionality, Revisited

The results pertaining to non-anthropocentrism may also be interpreted as suggesting that those six items represent two dimensions (Table 11). This observation could seem concerning because IRT assumes that items in a scale are unidimensional (i.e., represent a single underlying construct). However, the assumption of unidimensionality is routinely violated, due to factors unrelated to the underlying construct, such as respondents' attention, reader comprehension, and subject familiarity.

Nevertheless, one might think it valuable to formally assess dimensionality by performing CFA on the survey data, using the 13 concepts listed in Section 3.2 as a null hypothesis. Doing so may have been appropriate if the purpose of this research had been to use a general sample of non-experts as a baseline for evaluating the quality of how scholars have conceptualized issues in environmental philosophy. But that was not the purpose of the paper. Furthermore, there are important circumstances where limitations to CFA [87] make EFA preferable [88].

More importantly, EFA provided strong indication that respondents organize their beliefs about the environment in a manner that deviates considerably from the 13 concepts (Sections 5.2 and 6.3). With that insight, there seems little value in also performing CFA or giving further attention to the prospect that *non-anthropocentrism* appears to violate the assumption of unidimensionality, given the purpose of this paper. Recall that the primary purpose of this paper is to give due attention to content validity and the breadth of concepts in environmental thought. We used IRT in a way that was supportive to achieving that primary purpose. It is no less important that we encourage future research that refines measures of non-anthropocentrism.

6.9. Conceptual Nuance

Some important concepts in environmental scholarship are nuanced to the point of likely being a pernicious challenge for measurement among many members of a general public. Consider, for example, *Animism* and the items we considered to represent that concept, i.e., items about whether rocks, air, and water are alive (Table 12). We crafted those items—as best we could—to focus attention on each of those subjects. For example, one item was: "Is the water that you could pour into a glass alive?". Enough people responded to all three items affirmatively, such that the set of items had a narrow range of threshold values. A conventional reaction to that observation is an interest in developing items that perform better at the top and bottom ends of the scale. To that end, a colleague who reviewed this manuscript suggested considering items whose subject is different forms of water, such as rivers or oceans. Doing so may increase the range of threshold values, but those items would also introduce concerns about content validity. For example, "ocean" might be interpreted by some survey respondents as being no more than a large quantity of water, but others are likely to consider an ocean to be the water and the ecological community that lives in an ocean. For those survey respondents, the survey item is less representative of animism and more representative of holism, though not a great representation of holism.

These considerations about *Animism* (and *Holism*) lead to two salient points. First, this is an example of how a survey instrument can be subject to a vicious trade-off between maximizing content validity at the expense of other desirable statistical properties. Second, some important ideas in environmental scholarship may be too nuanced to be readily assessed among individuals who have little experience with those ideas.

6.10. Item Style

Some variation in responses to survey items can rise from surveys when they include items that are reverse-coded or items that differ with respect to their response sets (e.g., semantic differential v. agree/disagree). When an analyst's top priority is to maximize inter-item correlations, then it can be important to minimize the aforementioned sources of variation. We fully appreciate the good reasons for taking such an approach. But doing so may come at the cost of measurements that are artificially precise.

For example, we recommend including these items to represent *Fragility:*

F4—Nature tends to be resilient to human impacts.
F8—Nature tends to recover quickly from harm caused by humans.

Responses to F4 would be reverse-coded. If a survey respondent really does believe that nature is fragile to human impacts, then they should be able to answer both items consistently, even though one is reverse-coded, so long as both items have good content validity. If a respondent answers the items inconsistently (perhaps as a result of being unduly influenced by agreement bias), then the result would still be an appropriate measurement, i.e., a measurement that reflects the respondent's equivocalness.

These observations suggest another potential trade-off in survey design. In this case, the trade-off is between (1) minimizing inter-item correlation by standardizing the design of survey items and (2) allowing variation in the design of survey items for the sake of more accurate measurement.

6.11. Conclusions and Future Developments

By focusing on content validity and by trading the assumptions of classical test theory for those of item response theory, we generated highly interpretable results about the relative importance of various environmental beliefs for predicting environmental behaviors, behavioral intentions, and two overarching attitudes about the environment. In particular, *Sacredness* and *Hope* seem to be especially important predictors.

It is no less important that the pretext of this research is that most survey instruments for environmental beliefs were developed with sharp focus on statistical properties, such as factorial structure and correlations among items within hypothesized dimensions. That focus came at the expense of giving due attention to content validity. Some scholars

have suggested that instrument development may often involve an inescapable trade-off between content validity and forms of validity associated with factorial structure and inter-item correlations (e.g., [31]). This trade-off is concerning because an instrument with great statistical properties but low content validity is like having a scale that produces excellent measurements of an unknown property—or at least an ambiguously defined property. A central purpose of this paper has been to explore what instrument development looks when like the primary focus is content validity for an appropriately broad set of environmental concepts.

Because the assessment of content validity is inherently qualitative, it routinely depends on judgments. In that regard, assessments of content validity are provisional in the sense of being subject to insights from subsequent qualitative analysis. Future research with the survey items developed here should include the following: (1) testing different populations (e.g., outside the U.S. and different cultures within the U.S.); (2) developing items with broader ranges of threshold parameters for certain concepts; (3) developing concepts about the environment that we overlooked. Such research is warranted insomuch as progress toward sustainability depends on understanding how beliefs about the environment are related to behaviors that affect sustainability. Understanding those sustainability relationships can be greatly aided by survey instruments whose measurement is robustly interpretable. That interpretability depends on giving due attention to content validity.

Author Contributions: Conceptualization, J.A.V. and J.T.B.; development of survey items, all authors; formal analysis, J.A.V., B.G., C.E.R. and K.M.S.; writing—original draft preparation, J.A.V.; writing—review and editing, all authors; funding acquisition, J.A.V. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported, in part, by a grant to J.A.V. from the Templeton Religion Trust and the Issachar Fund.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Michigan Technological University (protocol code 1925680-1).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data upon which this research is based is available from J.A.V. upon request.

Acknowledgments: The authors are grateful to the slate of experts who reviewed the concepts and items developed through this research.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

PEB	pro-environmental behaviors
PEBI	pro-environmental behavioral intentions
IRT	item response theory
NEP	revised New Environmental Paradigm
CFA	confirmatory factor analysis
EFA	exploratory factor analysis
CHT	cognitive hierarchy theory
VBN	value-beliefs-norm theory
NAT	norm activation theory

Appendix A. Hypothesized Dimensionality of NEP and Revised NEP

All that was documented about the development of the hypothesized dimensions of NEP and the revised NEP is presented here. Specifically, NEP is reported to have been built by this method [18]:

> "The NEP items were carefully constructed by the researchers. First, in an effort to achieve content validity (Nunnally, 1967:79–83), we attempted to include items reflecting all of the crucial aspects of the NEP: limits to growth, balance of nature, anti-anthropocentrism, etc. In obtaining a representative set of items we were guided by our reading of the NEP literature cited above, and also consulted several environmental scientists at our university. The advice of the latter individuals was helpful in selecting a representative set of items and in wording the various items in an appropriate fashion."

The phrase, "NEP literature cited above", from the preceding passage is represented by this passage of text from [18]:

> "Numerous writers have argued that our nation's ecological problems stem in large part from the traditional values, attitudes and beliefs prevalent within our society (e.g., Disch, 1970). For example, it is often suggested that our belief in abundance and progress, our devotion to growth and prosperity, our faith in science and technology, and our commitment to a laissez-faire economy, limited governmental planning and private property rights all contribute to environmental degradation and/or hinder efforts to improve the quality of the environment (see, e.g., Caldwell, 1970; Campbell and Wade, 1972; Dunlap, 1976; Whisenhunt, 1974). Pirages and Ehrlich (1974:43–44) have argued that such a constellation of values, attitudes and beliefs comprises our society's "Dominant Social Paradigm" (or DSP). A DSP constitutes a world view "through which individuals or, collectively, a society interpret the meaning of the external world . . . [and] . . . a mental image of social reality that guides expectations in a society." Not surprisingly, they further argue that our society's fundamentally anti-ecological DSP must be replaced by a more realistic world view if ecological catastrophe is to be avoided.

Despite the predominance of an anti-ecological DSP within our society, new ideas have emerged in recent years which represent a direct challenge to this DSP. For example, we increasingly hear of the inevitability of "limits to growth," the necessity of achieving a "steady-state" economy, the importance of preserving the "balance of nature," and the need to reject the anthropocentric notion that nature exists solely for human use (see, e.g., Barbour, 1973; Commoner, 1971; Daily, 1973; Meadows, et al., 1972). Taken together, such ideas comprise a world view—perhaps best captured by the "spaceship earth" metaphor—which differs dramatically from that provided by our DSP. In recognition of this fundamental contrast, we term the new world view the "New Environmental Paradigm" or NEP."

The revised NEP states [12]:

> "The notion of "human exemptionalism," or the idea that humans—unlike other species—are exempt from the constraints of nature (Dunlap & Catton, 1994), became prominent in the 1980s through the efforts of Julian Simon and other defenders of the DSP. In addition, the emergence of ozone depletion, climate change, and human-induced global environmental change in general suggested the importance of including items focusing on the likelihood of potentially catastrophic environmental changes or "ecocrises" besetting humankind."

Finally, what had been three facets of the NEP (i.e., "limits to growth, balance of nature, anti-anthropocentrism, etc.") are later described by Dunalp et al. [12] as being "humanity's ability to upset the balance of nature, the existence of limits to growth for human societies, and humanity's right to rule over the rest of nature." No account is given for whether differences in the two previous passages of text are of any significance.

Appendix B. Sample Properties

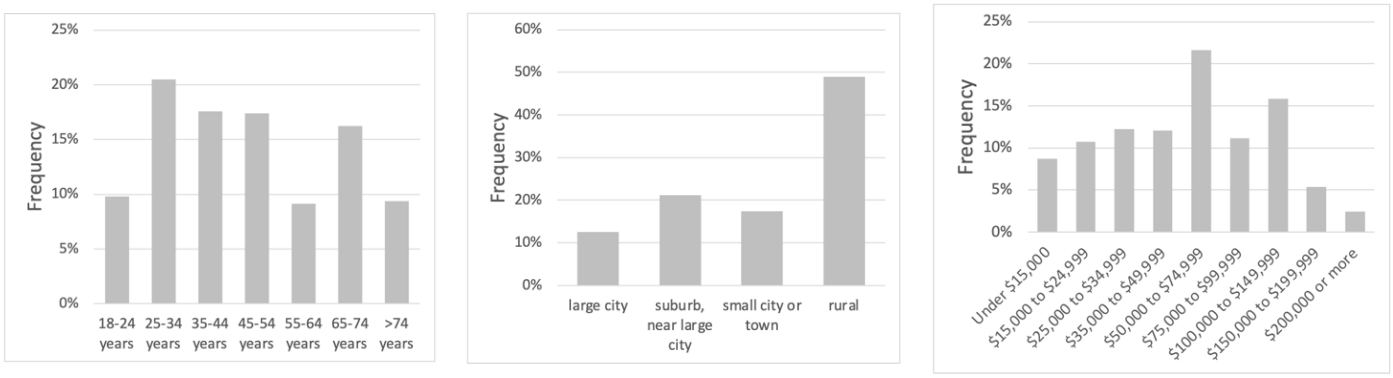


Figure A1. The distribution of age (left panel), living environments (middle panel) and income (**right panel**) for the sample of survey participants.

Compared to U.S. census data, 18–24 year-olds and 55–65 year-olds are slightly under-represented and those >65 are slightly over represented. Compared to census data, rural living people are significantly over represented in this sample. The sample is a reasonably close match to census data with respect to household income, except that the sample slightly underrepresents the two wealthiest income categories. Another pertinent sample property is that 51% of the participants indicated being female.

Appendix C. Scree Plot for Factor Analysis

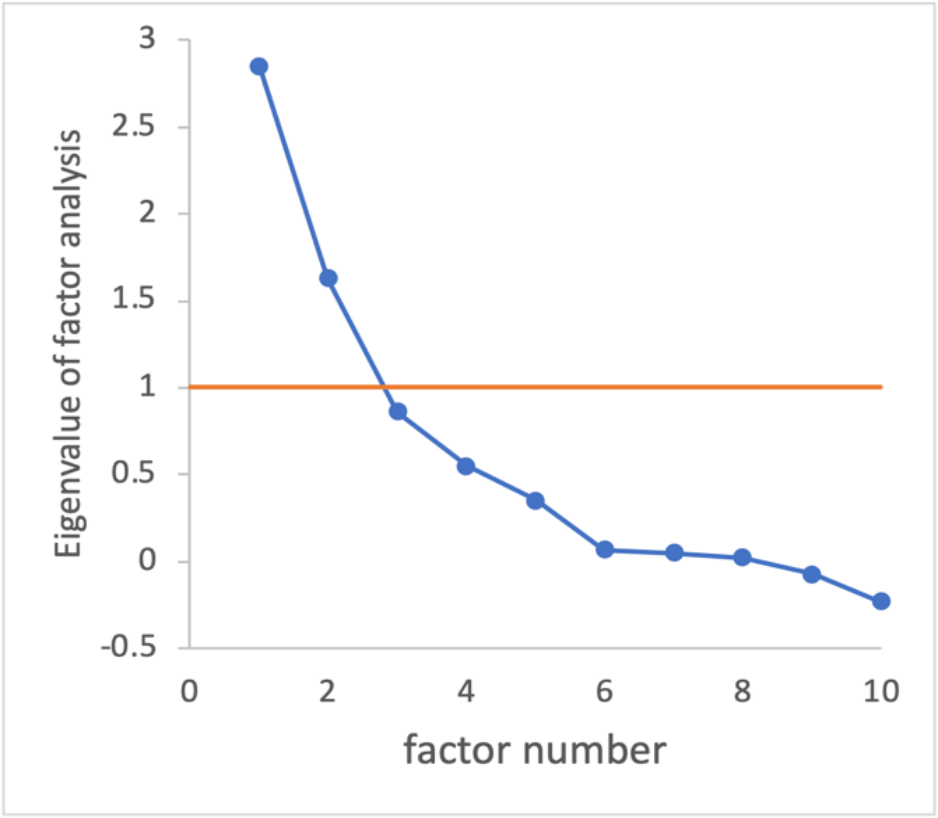


Figure A2. Scree plot for the factor analysis described in Section 5.2 of the manuscript.

Appendix D. Distribution of Scores for Pro-Environmental Behaviors and Behavioral Intentions

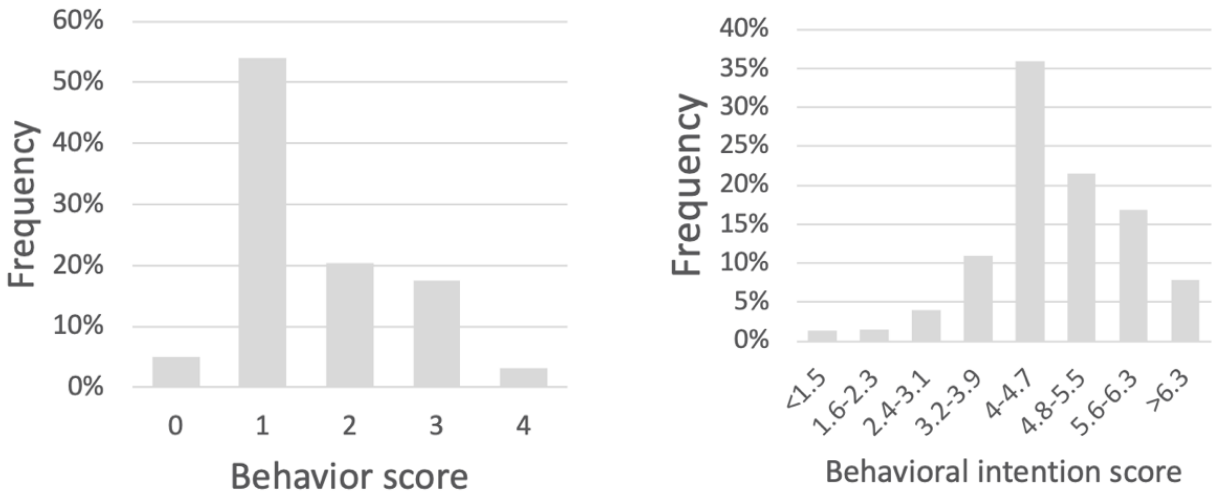


Figure A3. The distribution of scores for pro-environmental behaviors (**left panel**) and behavioral intentions (**right panel**).

The survey items from which these scores are based are given in Table 16. For survey items representing behaviors, a large portion of participants reported "never" or reported donating no money. For this reason, each participant's response to each item was scored as 0 or 1, and their behavior score is the sum of those zeros and ones. Each participant's score for behavioral intentions was their average response to the five items representing average responses to all 5-items.

Appendix E.

Table A1. Matrix of bivariate correlations (and standard errors).

	Dependency	Sacredness	Doubt	Stability	Holism	Animism	Non-Anthropocen1	Non-Anthropocen2	Comfort	Fragility	Connectedness	Shared Traits	Hope	Nature's Breadth
dependency	1													
sacredness	0.03 (0.05)	1												
doubt	−0.01 (0.05)	0.38 (0.04)	1											
stability	−0.17 (0.05)	0.04 (0.05)	−0.03 (0.05)	1										
holism	0.07 (0.05)	0.34 (0.04)	0.22 (0.05)	0.07 (0.05)	1									
animism	0.12 (0.05)	−0.30 (0.04)	−0.2 (0.05)	−0.07 (0.05)	−0.36 (0.04)	1								
Non-anthropocen1	0.52 (0.03)	0.18 (0.05)	0.12 (0.05)	−0.13 (0.05)	0.14 (0.05)	0.03 (0.05)	1							

Table A1. *Cont.*

	Dependency	Sacredness	Doubt	Stability	Holism	Animism	Non-Anthropocen1	Non-Anthropocen2	Comfort	Fragility	Connectedness	Shared Traits	Hope	Nature's Breadth
Non-anthropocen2	0.60 (0.03)	−0.05 (0.05)	−0.03 (0.05)	−0.17 (0.05)	0.00 (0.05)	0.10 (0.05)	0.56 (0.03)	1						
comfort	0.25 (0.04)	0.43 (0.04)	0.12 (0.05)	0.03 (0.05)	0.25 (0.04)	−0.06 (0.05)	0.30 (0.04)	0.2 (0.)	1					
fragility	0.58 (0.03)	0.18 (0.05)	0.18 (0.05)	−0.07 (0.05)	0.17 (0.05)	0.00 (0.05)	0.44 (0.04)	0.43 (0.)	0.25 (0.)	1				
connectedness	0.16 (0.05)	0.52 (0.03)	0.40 (0.04)	0.00 (0.05)	0.28 (0.04)	−0.18 (0.05)	0.25 (0.04)	0.03 (0.05)	0.29 (0.04)	0.1 (0.05)	1			
shared traits	0.16 (0.05)	0.36 (0.04)	0.26 (0.04)	−0.03 (0.05)	0.26 (0.04)	−0.22 (0.05)	0.30 (0.04)	0.15 (0.05)	0.22 (0.05)	0.27 (0.04)	0.34 (0.04)	1		
Hope	0.02 (0.05)	0.14 (0.05)	0.08 (0.05)	−0.03 (0.05)	0.15 (0.05)	−0.18 (0.05)	0.06 (0.05)	−0.04 (0.05)	−0.06 (0.05)	0.2 (0.05)	0.08 (0.05)	0.2 (0.05)	1	
nature's breadth	0.33 (0.04)	−0.17 (0.05)	−0.27 (0.04)	−0.12 (0.05)	−0.08 (0.05)	0.10 (0.05)	0.17 (0.05)	0.32 (0.04)	0.01 (0.05)	0.24 (0.04)	−0.22 (0.05)	0.01 (0.05)	0.04 (0.05)	1

References

1. Vucetich, J.A.; Nelson, M.P. Sustainability: Virtuous or vulgar? *BioScience* **2010**, *60*, 539–544. [CrossRef]
2. Cruz, S.M.; Manata, B. Measurement of environmental concern: A review and analysis. *Front. Psychol.* **2020**, *11*, 363. [CrossRef]
3. Cronbach, L.J.; Meehl, P.E. Construct validity in psychological tests. *Psychol. Bull.* **1955**, *52*, 281. [CrossRef]
4. Bhandari, P. Construct Validity | Definition, Types, & Examples. 2022. Available online: https://www.scribbr.com/methodology/construct-validity/ (accessed on 9 February 2024).
5. Dunlap, R.E.; Jones, R.E. Environmental concern: Conceptual and measurement issues. In *Handbook of Environmental Sociology*; Dunlap, R.E., Michelson, W., Eds.; Greenwood Press: Westport, CT, USA, 2002; pp. 482–524.
6. Heberlein, T.A. Environmental Attitudes. *Z. Fur Umweltpolit.* **1981**, *2*, 241–270.
7. Smythe, P.C.; Brook, R.C. Environmental concerns and actions: A social-psychological investigation. *Can. J. Behav. Sci. Rev. Can. Scie. Comp* **1980**, *12*, 175–186. [CrossRef]
8. Dunlap, R.E. The new environmental paradigm scale: From marginality to worldwide use. *J. Environ. Educ.* **2008**, *40*, 3–18. [CrossRef]
9. Arcury, T.A.; Christianson, E.H. Environmental worldview in response to environmental problems: Kentucky 1984 and 1988 compared. *Environ. Behav.* **1990**, *22*, 387–407. [CrossRef]
10. LaLonde, R.; Jackson, E.L. The new environmental paradigm scale: Has it outlived its usefulness? *J. Environ. Educ.* **2002**, *33*, 28–36. [CrossRef]
11. Bohlen, G.; Schlegelmilch, B.B.; Diamantopoulos, A. Measuring ecological concern: A multi-construct perspective. *J. Mark. Manag.* **1993**, *9*, 415–430. [CrossRef]
12. Dunlap, R.E.; Van Liere, K.D.; Mertig, A.G.; Jones, R.E. Measuring endorsement of the new ecological paradigm: A revised NEP scale. *J. Soc. Issues* **2000**, *56*, 425–442. [CrossRef]
13. Milfont, T.L.; Duckitt, J. The environmental attitudes inventory: A valid and reliable measure to assess the structure of environmental attitudes. *J. Environ. Psychol.* **2010**, *30*, 80–94. [CrossRef]
14. Milfont, T.L.; Gouveia, V.V. Time perspective and values: An exploratory study of their relations to environmental attitudes. *J. Environ. Psychol.* **2006**, *26*, 72–82. [CrossRef]
15. Rosa, C.D.; Fried, E.I.; Larson, L.R.; Collado, S. Four challenges for measurement in environmental psychology, and how to address them. *J. Environ. Psychol.* **2023**, *85*, 101940. [CrossRef]
16. Crocker, L.; Algina, J. *Introduction to Classical and Modern Test Theory*; Holt, Rinehart and Winston: Orlando, FL, USA, 1986.
17. Vucetich, J.A.; Bruskotter, J.T.; Nelson, M.P. Evaluating whether nature's intrinsic value is an axiom of or anathema to conservation. *Conserv. Biol.* **2015**, *29*, 321–332. [CrossRef] [PubMed]
18. Dunlap, R.E.; Van Liere, K.D. The "new environmental paradigm". *J. Environ. Educ.* **1978**, *9*, 10–19. [CrossRef]
19. Platt, J.R. Strong Inference. *Science* **1964**, *146*, 347–353. [CrossRef]

20. Maloney, M.P.; Ward, M.P.; Braucht, G.N. A revised scale for the measurement of ecological attitudes and knowledge. *Am. Psychol.* **1975**, *30*, 787–790. [CrossRef]
21. Weigel, R.; Weigel, J. Environmental concern: The development of a measure. *Environ. Behav.* **1978**, *10*, 3–15. [CrossRef]
22. Lounsbury, J.W.; Tornatzky, L.G. A scale for assessing attitudes toward environmental quality. *J. Soc. Psychol.* **1977**, *101*, 299–305. [CrossRef]
23. Schultz, P.W. The structure of environmental concern: Concern for self, other people, and the biosphere. *J. Environ. Psychol.* **2001**, *21*, 327–339. [CrossRef]
24. Dietz, T.; Fitzgerald, A.; Shwom, R. Environmental values. *Annu. Rev. Environ. Resour.* **2005**, *30*, 335–372. [CrossRef]
25. Stern, P.C.; Dietz, T. The value basis of environmental concern. *J. Soc. Issues* **1994**, *50*, 65–84. [CrossRef]
26. Schwartz, S.H.; Howard, J.A. A normative decision-making model of altruism. In *Altruism and Helping Behavior*; Rushton, J.P., Sorrentino, R.M., Eds.; Lawerence Erlbaum: Hillsdale, NJ, USA, 1981; pp. 89–211.
27. Edwards, M.C. An introduction to item response theory using the need for cognition scale. *Soc. Personal. Psychol. Compass* **2009**, *3*, 507–529. [CrossRef]
28. Toland, M.D. Practical guide to conducting an item response theory analysis. *J. Early Adolesc.* **2014**, *34*, 120–151. [CrossRef]
29. Walpole, H.D.; Wilson, R.S. A yardstick for danger: Developing a flexible and sensitive measure of risk perception. *Risk Anal.* **2021**, *41*, 2031–2045. [CrossRef]
30. Samejima, F. Estimation of latent ability using a response pattern of graded scores. *Psychom. Monogr. Suppl.* **1969**, *34*, 100. [CrossRef]
31. Danner, D.; Blasius, J.; Breyer, B.; Eifler, S.; Menold, N.; Paulhus, D.L.; Rammstedt, B.; Roberts, R.D.; Schmitt, M.; Ziegler, M. Current challenges, new developments, and future directions in scale construction. *Eur. J. Psychol. Assess.* **2016**. [CrossRef]
32. Eubank, B.H.; Mohtadi, N.G.; Lafave, M.R.; Wiley, J.P.; Bois, A.J.; Boorman, R.S.; Sheps, D.M. Using the modified Delphi method to establish clinical consensus for the diagnosis and treatment of patients with rotator cuff pathology. *BMC Med. Res. Methodol.* **2016**, *16*, 56. [CrossRef] [PubMed]
33. Gustafson, D.H.; Shukla, R.K.; Delbecq, A.; Walster, G.W. A comparative study of differences in subjective likelihood estimates made by individuals, interacting groups, delphi groups, and nominal groups. *Organ Behav. Hum. Perf.* **1973**, *9*, 280–291. [CrossRef]
34. Damasio, A.R. *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*; Houghton Mifflin Harcourt: Boston, MA, USA, 2003.
35. Naqvi, N.; Shiv, B.; Bechara, A. The role of emotion in decision making: A cognitive neuroscience perspective. *Curr. Dir. Psychol. Sci.* **2006**, *15*, 260–264. [CrossRef]
36. Koivisto, M.; Grassini, S. Affective responses to urban but not to natural scenes depend on inter-individual differences in childhood nature exposure. *J. Environ. Psychol.* **2022**, *82*, 101840. [CrossRef]
37. Ulrich, R.S. Aesthetic and affective response to natural environment. In *Behavior and the Natural Environment*; Springer: Berlin/Heidelberg, Germany, 1983; pp. 85–125.
38. Ducarme, F.; Couvet, D. What does 'nature'mean? *Palgrave Commun.* **2020**, *6*, 1–8. [CrossRef]
39. Vucetich, J.A. *Restoring the Balance*; Johns Hopkins University Press: Baltimore, MD, USA, 2021.
40. Neumayer, E. *Weak Versus Strong Sustainability: Exploring the Limits of Two Opposing Paradigms*; Edward Elgar: Northampton, MA, USA, 2003.
41. Egerton, F.N. Changing concepts of the balance of nature. *Q. Rev. Biol.* **1973**, *48*, 322–350. [CrossRef]
42. Kricher, J. *The Balance of Nature*; Princeton University Press: Princeton, NJ, USA, 2009.
43. McCann, K.; Hastings, A.; Huxel, G.R. Weak trophic interactions and the balance of nature. *Nature* **1998**, *395*, 794–798. [CrossRef]
44. Pimm, S.L. *The Balance of Nature? Ecological Issues in the Conservation of Species and Communities*; University of Chicago Press: Chicago, IL, USA, 1991.
45. Simberloff, D. The "balance of nature"—Evolution of a panchreston. *PLoS Biol.* **2014**, *12*, e1001963. [CrossRef]
46. Alarcon, G.M.; Bowling, N.A.; Khazon, S. Great expectations: A meta-analytic examination of optimism and hope. *Personal Individ. Differ.* **2013**, *54*, 821–827. [CrossRef]
47. Goodall, J.; Berman, P. *Reason for Hope: A Spiritual Journey*; Grand Central Publishing: New York, NY, USA, 1999.
48. Ojala, M. Hope and climate change: The importance of hope for environmental engagement among young people. *Environ. Educ. Res.* **2012**, *18*, 625–642. [CrossRef]
49. Wynn, G. Finding Hope: Environmentalism and the Anthropocene. *Hist. Geogr.* **2020**, *48*, 1–39. [CrossRef]
50. Lam, S.P. Predicting intention to save water: Theory of planned behavior, response efficacy, vulnerability, and perceived efficiency of alternative solutions 1. *J. Appl. Soc. Psychol.* **2006**, *36*, 2803–2824. [CrossRef]
51. Ostrom, E. Coping with tragedies of the commons. *Annu. Rev. Political Sci.* **1999**, *2*, 493–535. [CrossRef]
52. Morrell, K.; Dahlmann, F. Aristotle in the Anthropocene: The comparative benefits of Aristotelian virtue ethics over Utilitarianism and deontology. *Anthr. Rev.* **2022**, 20530196221105093. [CrossRef]
53. Gatta, J. *Making Nature Sacred: Literature, Religion, and Environment in America from the Puritans to the Present*; Oxford University Press: Oxford, UK, 2004.
54. Gottlieb, R.S. *This Sacred Earth: Religion, Nature, Environment*; Routledge: London, UK, 2003.
55. Hengst-Ehrhart, Y.; Schraml, U. Attitudes and Behaviors of Neopagans Toward Nature. *Ecopsychology* **2013**, *5*, 255–264. [CrossRef]
56. Milton, K. Nature is already sacred. *Environ. Values* **1999**, *8*, 437–449. [CrossRef]
57. Stone, J.A. *Sacred Nature: The Environmental Potential of Religious Naturalism*; Routledge: London, UK, 2017.

58. Keller, D.R.; Golley, F.B. Holism. In *The Philosophy of Ecology: From Science to Synthesis*; University of Georgia Press: Athens, GA, USA, 2000.
59. Allen, D.L. *Wolves of Minong: Their Vital Role in a Wild Community*; Houghton Mifflin: Boston, MA, USA, 1979.
60. Leopold, A.L. *Sand County Almanac: And Sketches Here and There*; Oxford University Press: Oxford, UK, 1949.
61. Lovelock, J.E.; Margulis, L. Atmospheric homeostasis by and for the biosphere: The Gaia hypothesis. *Tellus. Ser. A. Stockh. Int. Meteorol. Inst.* **1974**, *26*, 2–10. [CrossRef]
62. Odum, E. Pleasants. In *Fundamentals of Ecology*; Saunders: Philadelphia, PA, USA, 1967.
63. Pound, R.; Clements, F.E. *Phytogeography of Nebraska: I. General Survey*; Seminar; The Botanical Research Institute of Texas, Inc.: Lincoln, NE, USA, 1900.
64. Næss, A. A defense of the deep ecology movement. *Environ. Ethics* **1984**, *6*, 265–270. [CrossRef]
65. Quijada, J.B. 'Animism' recognizes how animals, places and plants have power over humans—And it's finding renewed interest around the world. *The Conversation*, 13 October 2022.
66. Waytz, A.; Cacioppo, J.; Epley, N. Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspect. Psychol. Sci.* **2010**, *5*, 219–232. [CrossRef]
67. Bird-David, N. "Animism" revisited: Personhood, environment, and relational epistemology. *Curr. Anthropol.* **1999**, *40*, S67–S91. [CrossRef]
68. Taylor, B.R. *Dark Green Religion: Nature Spirituality and the Planetary Future*; University of California Press: Berkeley, CA, USA, 2010.
69. Copi, I.; Cohen, C.; Rodych, V. *Introduction to Logic*; Routledge: London, UK, 2016.
70. Manfredo, M.J.; Urquiza-Haas, E.G.; Carlos, A.W.D.; Bruskotter, J.T.; Dietsch, A.M. How anthropomorphism is changing the social context of modern wildlife conservation. *Biol. Conserv.* **2020**, *241*, 108297. [CrossRef]
71. Preston, S.D.; De Waal, F.B. Empathy: Its ultimate and proximate bases. *Behav. Brain Sci.* **2002**, *25*, 1–20. [CrossRef] [PubMed]
72. DeVellis, R.F. Classical test theory. *Med. Care* **2006**, *44*, S50–S59. [CrossRef]
73. Baker, F. *The Basics of Item Response Theory*; ERIC Clearinghouse on Assessment and Evaluation: Washington, DC, USA, 2001.
74. Barrett, G.V.; Phillips, J.S.; Alexander, R.A. Concurrent and predictive validity designs: A critical reanalysis. *J. Appl. Psychol.* **1981**, *66*, 1–6. [CrossRef]
75. Dietz, T.; Gardner, G.T.; Gilligan, J.; Stern, P.C.; Vandenbergh, M.P. Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18452–18456. [CrossRef] [PubMed]
76. Cordano, M.; Welcomer, S.A.; Scherer, R.F. An analysis of the predictive validity of the new ecological paradigm scale. *J. Environ. Educ.* **2003**, *34*, 22–28. [CrossRef]
77. McFarlane, B.L.; Boxall, P.C. The role of social psychological and social structural variables in environmental activism: An example of the forest sector. *J. Environ. Psychol.* **2003**, *23*, 79–87. [CrossRef]
78. Kelsey, E. *Hope Matters: Why Changing the Way We Think Is Critical to Solving the Environmental Crisis*; Greystone Books Ltd.: Vancouver, BC, USA, 2020.
79. Ojala, M. Recycling and ambivalence: Quantitative and qualitative analyses of household recycling among young adults. *Environ. Behav.* **2008**, *40*, 777–797. [CrossRef]
80. Stern, P.C. New environmental theories: Toward a coherent theory of environmentally significant behavior. *J. Soc. Issues* **2000**, *56*, 407–424. [CrossRef]
81. Homer, P.M.; Kahle, L.R. A structural equation test of the value-attitude-behavior hierarchy. *J. Personal. Soc. Psychol.* **1988**, *54*, 638. [CrossRef]
82. Kuhn, T.S. The Function of Measurement in Modern Physical Sciences. *Isis* **1961**, *52*, 161–193. [CrossRef]
83. Tal, E. Measurement in Science. In *The Stanford Encyclopedia of Philosophy*; Zalta, E.N., Ed.; Stanford University Press: Stanford, CA, USA, 2020; Available online: https://plato.stanford.edu/archives/fall2020/entries/measurement-science/ (accessed on 9 February 2024).
84. Franklin, A.; Anderson, M.; Brock, D.; Coleman, S.; Downing, J.; Gruvander, A.; Lilly, J.; Neal, J.; Peterson, D.; Price, M.; et al. Can a Theory-Laden Observation Test the Theory? *Br. J. Philos. Sci.* **1989**, *40*, 229–231. [CrossRef]
85. Borsboom, D. *Measuring the Mind: Conceptual Issues in Contemporary Psychometrics*; Cambridge University Press: Cambridge, UK, 2005.
86. Vessonen, E. Operationalism and Realism in Psychometrics. *Philos. Compass* **2019**, *14*, e12624. [CrossRef]
87. Schmitt, R.S. Current methodological considerations in exploratory and confirmatory factor analysis. *J. Psychoeduc. Assess.* **2011**, *29*, 304–321. [CrossRef]
88. Orçan, F. Exploratory and confirmatory factor analysis: Which one to use first? *J. Meas. Eval. Educ. Psychol.* **2018**, *9*, 414–421.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.