Measuring pro-environmental behavior: review and recommendations

Florian Lange & Siegfried Dewitte

Behavioral Engineering Research Group, KU Leuven, Leuven, Belgium

Correspondence:

Florian Lange

BEE – Behavioral Engineering Research Group

KU Leuven

Naamsestraat 69

3000 Leuven

Belgium

florian.lange@kuleuven.be

This manuscript has been accepted for publication in the *Journal of Environmental Psychology* on 22 April 2019 (DOI: 10.1016/j.jenvp.2019.04.009).

Acknowledgements

This project received funding from the FWO and European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665501.

Abstract

Any scientific attempt to understand, predict, or promote pro-environmental behavior requires an adequate measurement tool for the assessment of pro-environmental behavior. The multidisciplinary interest in pro-environmental behavior has generated a large variety of such tools, ranging from domain-general and domain-specific self-report measures, field observations conducted with the help of informants, trained observers, or technical devices, to behavioral tasks for use in the laboratory. The present review discusses this broad spectrum of existing approaches to the measurement of pro-environmental behavior, their strengths and weaknesses, as well as possibilities to improve upon them. From this review, we deduce several recommendations for the development, selection, and application of measures in pro-environmental behavior research. We conclude by stressing the importance of established and validated measures for a cumulative science of pro-environmental behavior.

Keywords: pro-environmental behavior; conservation (ecological behavior); measurement; self-report; field observation; laboratory

1. Introduction

Human behavior is commonly accepted as a major contributor to various environmental issues including climate change, environmental pollution, and the loss of biodiversity (Stern, 1992; Swim, Clayton, & Howard, 2011; Wilson, 1988; Wynes & Nicholas, 2017). Addressing these issues requires understanding those human behaviors that mitigate or exacerbate them. This class of behavior has been examined under a plethora of different names (Larson, Stedman, Cooper, & Decker, 2015) in multiple fields of the behavioral sciences including environmental psychology (Steg & Vlek, 2009), organizational psychology (Norton, Parker, Zacher, & Ashkanasy, 2015), behavior analysis (Lehman & Geller, 2004), environmental education (Kollmuss & Agyeman, 2002), and consumer research (Peattie, 2010). Throughout this review, we will refer to this class of behavior as pro-environmental behavior (PEB), noting that it includes the commission of acts that benefit the natural environment (e.g., recycling) and the omission of acts that harm it (e.g., avoid air travel).

A crucial prerequisite for a scientific analysis of PEB is the ability to measure PEB. Underlying mechanisms and psychological correlates of a particular behavior can only be uncovered if this behavior can be accurately assessed. Similarly, the effectiveness of interventions to promote PEB can only be evaluated if assessment of the target behavior is possible. The present review discusses approaches to measuring PEB, their strengths and weaknesses, as well as possibilities to improve upon them.

As measures of PEB, we consider all attempts to quantify observable properties (i.e., frequency, latency, temporal extent, or intensity) of behaviors that impact the natural environment. Critically, research traditions differ in how they interpret these measures of behavioral properties (Nelson & Hayes, 1979) or in what they consider the actual object of measurement (Yoder, Lloyd, & Symons, 2018). Measured behavioral properties can be

viewed as context-specific characteristics of behavior itself or as indicators of latent characteristics of the behaving person. Both approaches require the initial quantification of behavioral properties (e.g., the assessment of recycling frequency or of the time spent under a hot shower) to be accurate. If an individual recycled a higher proportion of paper waste in 2018 than in 2017, a good measure of paper waste recycling should not take a higher value in 2017 than in 2018. Obtaining accurate information on observable properties of PEB is not a trivial task and much of this review is dedicated to scrutinizing whether existing measurement techniques are successful in addressing it. When PEB is viewed as an indicator of a person characteristic (rather than as a context-dependent sample of responding), measurement challenges are not limited to the accurate quantification of observable behavioral properties. In this case, inferences are made regarding the underlying latent construct that require additional assumptions (e.g., some degree of stability across contexts). As detailed in sections 2.1 and 5.1, these inferences and assumptions are particularly relevant for research on individual differences in the propensity to engage in PEB. In contrast, less additional assumptions are involved when context-specific PEB characteristics are considered to be relevant in and of themselves (e.g., when they are assessed as outcome measures in experimental studies, see also sections 5.2 and 5.3). In practice, objects of behavioral measurement lie on a continuum between two extremes (i.e., context-dependent responding and generalized person characteristic), differing in the level of inference involved in interpreting quantifications of behavioral properties (Yoder et al., 2018).

An additional note seems warranted with regard to the selection of the behaviors reviewed here. As indicated in the definition above, we followed an impact-oriented approach in focusing on behaviors that actually affect the natural environment. Intent-oriented measures (e.g., the self-report item “Have you ever taken any action out of concern for climate change?”, Whitmarsh, 2009) do not necessarily assess behavior with actual environmental

impact and are not further discussed in our review. Moreover, while all types of PEB, by definition, involve positive consequences for the environment as a common denominator, classes of PEB also seem to differ substantially from each other (Stern, 2000; Truelove & Gillis, 2018). For example, many categorizations of PEBs involve a distinction between private-sphere and public-sphere behaviors (Larson et al., 2015; Stern, 2000; Stern, Dietz, Abel, Guagnano, & Kalof, 1999). While we did not set out to exclusively review measures of private-sphere PEBs, the reader will see that most of the measures and examples discussed in this article pertain to conservation behaviors that occur within the private sphere. To some extent, this focus is reflective of the field of PEB research, as the variety of approaches to study public-sphere PEB has been limited thus far. For example, we have no knowledge of approaches to quantify environmental activism by objectively observing behavior in the field or laboratory. Given this focus on private-sphere PEB, readers should keep in mind that not all of the conclusions made in this review might be generalizable to all PEB domains.

This review is intended to be of practical use to everyone who wishes to measure PEB. To this end, we will first provide an illustrative overview of the wide spectrum of existing measurement approaches. Starting with the review of self-report measures, we will go on to discuss field observation methods, before turning to the laboratory assessment of PEB. We will conclude by deducing recommendations on how to select a measurement approach given a particular research question.

2. Self-report measures of pro-environmental behavior

Self-report assessment entails that individuals are asked to provide information on the properties of the behaviors they perform in everyday life. Individuals can respond to this request in the course of interviews, via (e)mail, or by completing online questionnaires. Self-report data can typically be collected at a low cost, which makes self-report PEB measures

attractive to researchers requiring large sample sizes or for inclusion into large-scale (international) social survey research (e.g., Pisano & Lubell, 2017; Tam & Chan, 2017).

Self-report assessment can target different behavioral properties by asking individuals, for example, if they engage in a PEB at all, how frequently they engage in it, or how pro-environmental a particular behavior of theirs is (e.g., *At which temperature do you wash your clothes?*). Questions can refer to different time frames, ranging from the present (e.g., *How often do you…?*) to a specified (e.g., the past month or year) or unspecified interval in the past. In addition, self-report measures of PEB differ with regard to their specificity (see Vining & Ebreo, 2002, for a discussion of the relevance of this dimension). Items can either refer to PEB in general (e.g., "*I participate in pro-environment behaviors.*", Obery & Bangert, 2017) or specify the characteristics of a particular PEB in question (e.g., "*In the past month, when I am at home I recycle paper.*", Maki & Rothman, 2017).

While some authors rely on single-item measures to assess specific or general PEB, others construct more comprehensive multi-item scales that are typically less affected by measurement error (Churchill, 1979). Multi-item scales for the assessment of PEB are highly diverse. For example, Markle (2013) identified not less than 42 unique multi-item PEB measures in 49 reviewed studies. Many of these scales are *ad hoc* measures of unknown psychometric quality that have been developed for a particular research project (Dono, Webb, & Richardson, 2010).

Other researchers create PEB scales based on an explicit psychometric analysis of item and scale properties. This practice provides others with the evidence-based confidence necessary to use the same validated scale in their own study, thus contributing to a cumulative science of PEB. An overview (Table S1) and discussion of established multi-item self-report measures of PEB can be found in the Supplementary Materials. A large number of these scales has been designed for the assessment of an individual's propensity to engage in pro-

environmental behavior across different domains. Based on its frequency of use and thoroughness of psychometric evaluation, the General Ecological Behavior (GEB) measure (Kaiser, 1998; Kaiser & Wilson, 2004) can probably be considered the best established of these domain-general propensity measures. Next to such global PEB measures, more specific scales exist that focus on particular populations (e.g., children, Evans et al., 2007; Kaiser et al., 2007), particular contexts (e.g., the workplace, Boiral & Paillé, 2012; Robertson & Barling, 2017), or particular domains of PEB (e.g., activism, Alisat & Riemer, 2015; or consumption behavior, Roberts, 1996).

Further adding to the toolbox of the PEB researcher, diary procedures differ from the scales described above in that they require participants to report their behavior on multiple occasions. Self-report diaries have been used, for example, to have participants indicate, for every day, how many items of paper they recycled (Chu & Chiu, 2003), when they switched on and off their office lights (Maleetipwan-Mattsson, Laike, & Johansson, 2013), or the characteristics (e.g., duration, distance, travel mode) of each trip they took (Bamberg, 2006).

Finally, self-reports of PEB are also used to create ecological footprint measures (Bleys, Defloor, van Ootegem, & Verhofstadt, 2018; Huddart Kennedy, Krahn, & Krogman, 2015). Rather than PEB per se, footprint measures assess the product of behavior and its environmental significance. To this end, participants are asked to report on a number of PEBs and the resulting data are then multiplied with the associated amounts of energy used or carbon emissions produced.

2.1 Limitations of self-report measures of pro-environmental behavior

The validity of self-report measures assessing properties of PEBs has often been questioned (Gifford, 2014; Lange, Steinke, & Dewitte, 2018). In order for a self-report item to qualify as a valid PEB measure, responses to this item need to correspond to the properties of the respective behavior. For example, if Person A recycles paper more often than Person B,

Person A should indicate a higher frequency of paper recycling in response to the question "*How often do you recycle paper*?" than Person B. Similarly, responses of Person A should scale with paper-recycling fluctuations in the everyday life of Person A. These assumptions, however, might not always be very realistic.

First, it is highly unlikely that all respondents have the same idea of the concepts of "paper", "recycling", and "often" (see also Kormos & Gifford, 2014). Even within individuals, the answer to the question of how often is "often" could change, for example, after an intervention. Second, such a question does not ask for a simple behavioral report, but rather for an extensive retrospective survey including appropriate aggregation procedures. Upon presentation of the item "*How often do you recycle paper*?", respondents are in a particularly bad position to do such a survey. Until a few seconds ago, they did not know that this was their task, they are not trained in behavioral observations, they may have forgotten many instances of paper recycling, and they might not be very motivated to spend large amounts of time to conduct a thorough survey for every single item. Repeated assessment of PEB (e.g., in the context of diary studies) or inquiring about dichotomized practices or circumstances (e.g., car ownership, Kaiser, Frick, & Stoll-Kleemann, 2001) may reduce these survey demands and increase the accuracy of self-reports. Third, individuals are not impartial observers of their own behavior. They may want their response to be consistent with the other responses they gave in the study, the way they would like to behave, or the expectations or preferences of the researcher. Studies examining the last possibility typically find small and non-significant correlations between self-report measures of PEB and social desirability scales (Milfont, 2009).

In sum, there are many factors that can compromise the validity of answers to questions like "*How often do you recycle paper?*". This does not imply that the scores produced on self-report measures of PEB are meaningless. Just like questions about

environmental attitudes and intentions, PEB items might reflect an individual's propensity to engage in PEB (Kaiser, Byrka, & Hartig, 2010). All other things being equal, an individual with a high propensity to engage in PEB can be expected to show more PEB than an individual with a low propensity. Identifying the determinants of such a propensity might be a fruitful endeavor, for example, for personality researchers (Brick & Lewis, 2016; Markowitz, Goldberg, Ashton, & Lee, 2012). However, information about pro-environmental propensities cannot be used to infer the actual frequency (or other properties) of PEB in everyday life, simply because all other things are not equal (i.e., behavior occurs in a dynamic context of costs, constraints, and competing propensities).

Finally, self-report measures are difficult to use in experimental PEB studies (Lange et al., 2018). When participants are asked to survey their everyday PEB themselves, they must be given a sufficient amount of time to change their behavior after exposure to an experimental manipulation. Self-reports of PEB thus have to be collected in multiple testing sessions, a requirement that might discourage many researchers. As a result, researchers interested in causal relationships often resort to self-report measures of antecedents of PEB (e.g., intention) or hypothetical scenarios and thus end up studying verbal behavior (rather than behavior with actual environmental consequences) that can be shown at no cost (see Klein & Hilbig, 2019, for data on the relevance of studying consequential behavior). An alternative to this approach is measuring PEB in its context, which will be discussed in the following sections.

3. Field observations of pro-environmental behavior

Field observations of PEB promise a certain degree of objectivity as they acquire information about behavioral properties without relying on the subjective report of the behaving individual. Similar to self-report measures, field observations of PEB can take many different forms. We follow Kormos and Gifford (2014) in distinguishing these approaches

based on how the observation is conducted: by informants, trained observers, or the use of device measurements.

3.1 Informant reports

Informant reports are obtained from well-acquainted others, such as friends, spouses, or co-workers of the target individual (Vazire, 2006). These informants are either asked for a retrospective account of their casual observations (e.g., Seebauer, Fleiß, & Schweighart, 2017) or to deliberately observe target individuals for a given time before reporting on their behavior (e.g., Lam & Cheng, 2002). Report forms closely resemble those used in self-report research on PEB. For example, Lam and Cheng (2002) derived their informant measure from a self-report measure by replacing, among others, the item "*Do you recycle paper?*" with the item "*Does your spouse recycle paper?*". Of course, such measures might be affected by problems similar to those affecting self-report measures. Informants might tend to produce observation records that are consistent with their view of how the target individual is or should be. To improve objectivity, informants can be trained and the agreement between multiple raters observing the same behavior (i.e., the inter-rater reliability) can be evaluated. This approach is exemplified by a study by Chao and Lam (2011) who ascertained that PEBs of dormitory residents were rated similarly by all of their roommates before using the roommates' observations for hypothesis-testing analyses.

3.2 Trained observers

Instead of relying on recruited informants, researchers can also conduct behavioral observations themselves or train students in observing selected PEBs. In contrast to informants, trained observers can focus exclusively on the task of observing the target individual's PEB. This concentration on the variable of interest and the lack of a personal relationship with the target can be expected to increase the accuracy of behavior recordings.

Observational methods differ with regard to the object of observation (behavior vs. products of behavior) and the observational context (naturalistic vs contrived situations; Kazdin, 1979).

PEB has been directly observed in naturalistic situations by registering the travel mode of traffic participants (Mayer & Geller, 1982-1983), counting the number of returnable bottles in the shopping cart of grocery shop customers (Geller, Farris, & Post, 1973), and recording whether car drivers turned off their engines at a closed level crossing (Meleady et al., 2017). In contrast, researchers analyzing the composition of participants' garbage (Corral-Verdugo, Bernache, Encinas, & Garibaldi, 1994-1995; Cote, 1984), recording whether machines were turned off when unoccupied (Siero, Bakker, Dekker, & van den Burg, 1996), counting recycling bins at the curb (Gamba & Oskamp, 1994), or checking for non-reused towels on the floor of hotel rooms (Goldstein, Cialdini, & Griskevicius, 2008) rather observed the products of PEB (also referred to as behavioral residue, Gosling, Ko, Mannarelli, & Morris, 2002). They did not witness individuals actually performing a PEB. The fine line between observing behavioral products or behavior itself can be illustrated by studies assessing whether individuals turned off the lights after exiting a room. While Bergquist and Nilsson (2016) directly observed the behavior of switching off the lights, Dwyer, Maki, and Rothman (2015) assessed the status of the lights before and after individuals entered the black box of a bathroom, and then inferred behavior from changes in light status (see also Murtagh, Gatersleben, Cowen, & Uzzell, 2015). To the degree to which it can be ensured that a chosen product can only be produced by the PEB of the target individual, observations of behaviors and behavioral products can be treated as practically equivalent.

When baseline frequencies of a PEB are low or when it is important that all individuals have similar opportunities to perform the behavior, it can be sensible to contrive a situation that facilitates the behavior of interest. Recording how much money visitors of a national park donate in response to a request for supporting the park (Alpizar, Carlsson, &

Joansson-Stenman, 2008) is a straightforward example of observing PEB in a contrived situation. Examples for observing behavioral products in contrived situations include assessing whether participants correctly dispose of a handbill distributed in a grocery shop (Geller, Wittmer, & Tuso, 1977), attach a previously distributed sticker prohibiting advertisements to their mailbox (Hamann, Reese, Seewald, & Loeschinger, 2015), or mail back a public transport ticket they received from the experimenter after having used it (Bamberg, 2002; see also Katzev & Bachman, 1982).

Observations of behavior in a naturalistic or contrived situation should be unobtrusive in order not to evoke reactance (Kazdin, 1979; 1982). For example, observers in the studies cited above were stationed in a parking car (Mayer & Geller, 1982-1983) or in an office having an unobstructed view on the scene of interest (Murtagh et al., 2015). This requirement is easier to meet when observing products of PEB that cannot be affected by the observer's presence. The use of camera recordings, in accordance with ethical standards, can help achieving comparable levels of unobtrusiveness for direct observations of PEB.

Errors and bias on part of the trained observer are commonly considered "not harmful unless they go undetected and unmeasured" (McCall, 1984, p. 273). Detection and measurement typically occurs in the course of evaluating inter-rater reliability. In addition, careful selection, training, and supervision of observers is required to ascertain that behavior ratings are not systematically distorted by observers' prejudices or expectations. Observer drift (i.e., "implicit changes in code definitions made by observers over time", Smith, 1986, p. 720) represents another possible threat to the validity of observer ratings. This risk can be controlled by testing coding systems for unambiguity and exhaustiveness before using them in a field study.

3.3 Device measurements

When observing PEB in the field, researchers can also draw on a variety of technical devices. Most often, these devices do not assess PEB directly but rather a PEB product. An early example is provided by Foxx and Hake (1977) and Hake and Zane (1981) who checked participants' odometers to calculate the distance travelled by car. Household consumption of electricity, gas, and water is another popular variable for device-mediated measurement. These data can be obtained by visiting participating households to monitor their utility meters (Schultz et al., 2016; Winett & Nietzel, 1975). Other researchers have requested consumption (Gregory & Di Leo, 2003) or billing data (Sapci & Considine, 2014) from utility companies or inspected participants' thermostat settings (Walker, 1979). With regard to such consumption data, the gap between particular PEBs and the observed behavioral product is particularly large. Meter readings are aggregate products of all utility-consuming behaviors of all individuals who have access to the utilities tracked by a particular meter. Hence, changes in meter readings cannot be attributed to a particular behavior of a particular individual (Gatersleben et al., 2002). This gap between meter readings and behavior could be bridged by tracking utility consumption of specific devices that are typically only used by one individual (e.g., smartphones). Alternatively, it is possible to use measurement devices to track PEB itself rather than its products. The extant measurement approach that comes closest to this idea might be the assessment of speeding behavior by the use of GPS technology installed in participants' cars (Bolderdijk, Knockaert, Steg, & Verhoef, 2011). GPS data might also allow for the identification of travel mode choices (Brown et al., 2016; Xiao, Juan, & Zhang, 2015), which would open interesting possibilities for future field studies on PEB.

3.4 Limitations of field observations of pro-environmental behavior

The main factor accounting for popularity differences between self-report measures and field observations of PEB might be data-collection cost. Measuring PEB in the field often requires more financial resources (e.g., for paying trained observers), time (e.g., to collect a

sufficiently large sample), and preparatory efforts (e.g., to obtain approval from an ethics committee) than distributing an online questionnaire. In addition, many of the questions that are typically addressed via self-reports are difficult to address with observational data. For example, examining the relationship between self-report measures of potential PEB predictors and field measurements of PEB would necessitate having individuals complete a questionnaire while, after, or before observing their behavior in an unobtrusive way. In many of such cases, measurement of the predictor variables might distort measurement of the outcome variable. Field observations appear more useful in experimental research (e.g., when evaluating the effectiveness of interventions to promote PEB). However, the complexity of field settings often undermines the validity of experimental field research. Researchers may not always have sufficient control over the experimental situation to randomly assign participants to different conditions. For example, if a study were to administer information about an environmental issue via posters or billboards, the information would be perceived by target individuals and their neighbors (who are thus ineligible for a no-intervention control group). If the same information were mailed to target individuals, they might talk about it with their neighbors, who might then wonder why they are treated differently. Hence, when studying such an intervention in the field, randomization cannot occur on the level of individuals, but only on higher levels (e.g., neighborhoods, Keller, 1991; cafeterias, Dupré & Meineri, 2016; or residence halls, Mallett & Melchiori, 2016). Moreover, experimenters may not always be able to reliably track all participants contributing data points to their observation. In the field studies by Murtagh and colleagues (2015) and Bergquist and Nilsson (2016), the unit of analysis were visits of a room (where individuals could turn off the lights or not). The same individual might have visited the room multiple times and thus contributed multiple observations in the same or different experimental conditions. Finally, the difficulties of collecting background data from observed individuals mentioned above further constrain

the possibilities associated with field experimental research. Field experiments on PEB typically involve neither manipulation checks nor the assessment of potentially relevant control variables (see Hamann et al., 2015, for discussion). Relatedly, it is often impossible to relate the effectiveness of interventions to individual-difference variables (which would be necessary for developing tailored interventions).

4. Laboratory observations of pro-environmental behavior

In contrast to field observations, a higher degree of experimental control can be achieved when studying PEB in the laboratory. In the laboratory, participants are per definition exposed to a contrived situation and conditions can be arranged for experimenters to directly observe a type of PEB as it unfolds. Some researchers have tried to inconspicuously embed an opportunity for showing PEB within a sequence of tasks given to the participant. For example, Murtagh and colleagues (2015) assessed whether participants turned off the laboratory lights before switching to another testing room. Similarly, Huffman, Van Der Werff, Henning, and Watrous-Rodriguez (2014) asked their participants to dispose of the materials they were given for a mock task. When doing this, participants could choose between a recycling bin and a trash bin and they were awarded one point for every material that was correctly disposed of. In a further laboratory study involving multiple measures of PEB, Cornelissen, Pandelaere, Warlop, and Dewitte (2008) assessed participants' choice between recycled and regular paper when offering them a notepad as a gift as well as how much scrap paper participants used whilst completing a mock task.

Other researchers exposed participants to explicit tasks on which they could behave pro-environmentally or not. Most often, such tasks involve the possibility to allocate money in a pro-environmental way. Participants in the study by Barber, Bishop, and Gruen (2014) were asked to use their participation fee to make a bid for organic vs. conventional wine in an auction task. Similarly, participants tested by Vesely and Klöckner (2018) earned money in

one part of the study and could then donate parts of it to an environmental organization of their choice. Such tasks can also be administered over the internet. Hanss and Böhm (2013) endowed the participants of their online study with a small budget they could use to purchase either organic or conventional products. Similarly, participants received small fees for participating in the online study by Clements, McCright, Dietz, and Marquart-Pyatt (2015) and they could donate parts of this budget to an environmental organization. In another online study, participants used a navigation system to choose between different travel routes (Taube, Kibbe, Vetter, Adler, & Kaiser, 2018). Routes were either associated with long waiting times and large emission savings (i.e., large donations, made by the researcher, to an environmental organization that compensates for greenhouse gases) or with short waiting times and small emission savings. Hence, choosing the pro-environmental route involved actual waiting-time cost for the participants. Despite its relative convenience, the online administration of behavioral tasks is also associated with a loss of experimental control when compared to assessment in the laboratory. Participants might not be focused on the online survey, consult outside sources, or forgo the consequences of their behavior (e.g., by pursuing alternative activities during the waiting periods in the task by Taube et al., 2018).

The laboratory tasks reviewed thus far are *ad hoc* measures of PEB. They have been used based on the rationale that the behaviors involved have obvious consequences for the environment and can thus be considered to be pro-environmental. Yet, the face validity of those tasks does not abolish the need for standardized and psychometrically evaluated measurement tools. Such tools would allow for assessing PEB at known levels of reliability and validity, thereby yielding results that can be meaningfully compared across studies and laboratories. In addition, they can be considered an antidote against researcher degrees of freedom (Simmons, Nelson, & Simonsohn, 2011). Where an established protocol for the collection and analysis of PEB data exists, there is less room for arbitrary methodological

choices that might lead to the inflation of false-positive rates and effect sizes. Despite these advantages, only few established laboratory tasks have found application in the field of PEB research.

One of these rare established tasks related to the assessment of PEB is the FISH simulation developed by Gifford and colleagues (Gifford & Gifford, 2000; Gifford & Wells, 1991). In this task, participants act as fishers deciding how many fish to catch across multiple seasons. For each fish they catch, participants receive a small amount of money. If some fish are left in the ocean at the end of a season, the resource (i.e., fish) can regenerate at a rate chosen by the experimenter. Critically, participants do not make fishing decisions in isolation, but play together with other participants or computer-simulated fishers. These conditions create a commons dilemma. Fishers can either maximize their personal short-term gain or restrain themselves with an eye on the sustainable long-term management of the common resource. A typical FISH outcome measure is the proportion of fish taken by an individual, which can be interpreted as an indicator of preservationist resource-management practices (Gifford & Hine, 1997). The task allows for user-defined changes of numerous parameters (e.g., the level of "greed" of the computer-simulated fishers). The most recent task manual can be found at http://web.uvic.ca/~esplab/?q=tools.

It should be noted that fishing behavior in this simulation does not have actual consequences for the environment (see Tarditi, Hahnel, Jeanmonod, Sander, & Brosch, 2018, for a recently developed but not yet explicitly validated social dilemma task with environmental consequences). Fishers impact the simulated environment, but this impact translates to the real world only as consequences for the fisher and any potential fellow players (who might have fewer fish to catch in following seasons, resulting in a smaller payout). Hence, the behavior of a restrained fisher might rather be considered to be long-term oriented, cooperative, and economically sustainable than truly pro-environmental.

In contrast, the recently developed Pro-Environmental Behavior Task (PEBT, Lange et al., 2018) involves actual consequences, not only for the participant, but also for the environment. The task requires participants to make a number of trips. For each trip, they can choose between an environmentally friendly (e.g., the bicycle) and an environmentally unfriendly (e.g., the car) mode of transportation. Following their choice, they have to endure a waiting period, which is typically longer for the environmentally friendly than for the environmentally unfriendly option. However, whenever participants choose the environmentally unfriendly option, a series of USB-powered lights is illuminated for the duration of the trip. The associated waste of energy and CO_2 emissions make this option truly environmentally unfriendly and choosing the environmentally friendly PEBT option an actual PEB. The proportion of trials on which participants choose the environmentally friendly option indicates how participants trade off personal and environmental consequences. It has thus been proposed to be a suitable measure of PEB in the laboratory, a notion that has received support in recent validation studies (Lange et al., 2018). The framing and parameters of the PEBT can easily be adapted to allow addressing particular research questions. Researchers interested in using the task can download it at https://osf.io/tcnza/.

Another laboratory task that has recently been developed and validated is the Greater Good Game (GGG, Klein & Hilbig, 2018; Klein, Hilbig, & Heck, 2017). The game is played by three participants, who all receive a small monetary endowment at the beginning of each trial. Participants can either keep this money to themselves, donate it to a group account, or donate it to an environment account. Donations to the group account are doubled by the experimenter and then equally distributed among all participants. Donations to the environment account are doubled as well and then donated to an environmental organization. This procedure is repeated multiple times and payoffs of a randomly selected trial are actually paid out to participants and the environmental organization. Due to its forced-choice task

structure, the task allows dissociating forms of cooperative behavior and PEB. Precise task instructions for the GGG can be found at https://osf.io/zw2ze/.

4.1 Limitations of laboratory observations of pro-environmental behavior

Laboratory studies are, per definition, conducted in an artificial environment. With this in view, some researchers consider studies using laboratory measures of PEB to lack ecological validity (e.g., Jacobs & Harms, 2014; Sörqvist, Haga, Holmgren, & Hansla, 2015). Ecological validity is an elusive construct that has been inconsistently defined (Bronfenbrenner, 1977; Schmuckler, 2001). For example, an investigation might be "regarded as ecologically valid if it is carried out in a naturalistic setting and involves objects and activities from everyday life" (Bronfenbrenner, 1977, p. 515).

When following such a definition, one can only conclude that research using the PEBT, for example, must be ecologically invalid. Everyday life choices between the car and the bicycle are not made by clicking on a symbol on a computer screen. The PEBT is not completed in a "naturalistic setting" (but in the laboratory), and it does not involve objects (e.g., cars and bicycles) nor activities (e.g., searching for keys) "from everyday life". Some researchers may fear that this artificiality of laboratory situations critically limits the generalizability of research conducted in the laboratory in general.

Fortunately, a closer look at issues of ecological validity and generalizability suggests that these concerns are largely unwarranted (Bronfenbrenner, 1977; Schmuckler, 2001). The definition given above is typically considered to be too simplistic and misleading (e.g., Bronfenbrenner, 1977; Dunlosky, Bottiroli, & Hartwig, 2009). Findings obtained in an artificial lab environment might be highly generalizable and findings obtained in a naturalistic field setting might not generalize beyond this specific setting at all. What, then, determines the generalizability of results beyond the context in which they have been produced?

According to Schmuckler (2001), "the issue involves identifying the critical theoretical parameters underlying the psychological processes in question and then determining whether these parameters occur in the empirical context" (p. 432). This definition illustrates the close link between the ecological validity of an investigation and its theoretical background. In the case of the PEBT, for example, the task was created to reflect the "conflict between individual and environmental consequences, which is characteristic of many environmentally significant decisions in everyday life" (Lange et al., 2018, p. 47). In the words of Schmuckler (2001), individual consequences, environmental consequences, and the conflict between them are identified as "the critical theoretical parameters" underlying pro-environmental decision-making. According to his definition, findings obtained with the PEBT in the laboratory can fail to generalize for two reasons.

First, the identified theoretical parameters might not be critical for PEB. Laboratory PEBT findings can only be expected to generalize to the extent that the analysis regarding the conflict underlying environmentally significant decisions is accurate. They are unlikely to generalize to behaviors that are primarily driven by other parameters. This implies that one would expect them to generalize to some everyday situations (i.e., those that primarily involve this conflict), but not to others (i.e., those that are dominated by a different conflict).

Second, the identified theoretical parameters might not occur in the empirical context. PEBT findings from the laboratory can only be expected to generalize to the extent that the task structure accurately reflects the conflict between individual and environmental consequences. One might argue, for example, that the task structure does not do so because the waiting times on the PEBT and the amount of energy consumed by the USB-powered PEBT lights are negligible. The question of whether the operationalizations of these parameters are effective in establishing a conflict between them is an empirical one that can and should be addressed in the context of validation studies (e.g., Lange et al., 2018).

The above analysis illustrates that laboratory measures of PEB can produce generalizable results (despite being artificial) if they involve an effective operationalization of the critical parameters underlying the PEB(s) of interest. Variables that can be shown to affect behavior on a laboratory task that meets these criteria can be expected to also affect those everyday pro-environmental decisions that are shaped by the same critical parameters. Notably, this does not guarantee that a field study on a PEB will find an effect of similar size when manipulating the variable in the same way as in the laboratory. Field situations involve a degree of noise that may render a small effect from the laboratory very difficult to detect. This does not imply that the effect is practically meaningless, but rather that the intervention has to be adjusted to exert appreciable effects in the field. We will revisit this issue in the following section.

5. Recommendations

The last decades of PEB research have produced a large diversity of measurement tools. Parts of this diversity can be attributed to the variety of research priorities and methodological preferences that emerges from the multidisciplinary interest in PEB. Other parts might rather reflect a tendency to create idiosyncratic *ad hoc* measures that seem to be best suited to address the research question at hand. This tendency is unfortunate as it stands in the way of a cumulative science of PEB. Note that it makes sense not to use an established measure just because of it being established when this measure does not meet the requirements of a particular research project. Similarly, it is likely that for some questions about PEB, searches for a suitable established measure will be in vein. However, it does not follow that researchers in such situations should use just any measure to assess PEB. If a research question cannot be answered convincingly because of the lack of a suitable established measure of PEB, it might be advisable to take a step back to systematically develop such a measure first.

For many research goals, however, a suitable established tool to measure PEB might already exist. This raises the question of how to identify this tool from the large number of measures reviewed in this paper. Unsurprisingly, the answer to this question depends on the objectives of the particular research project. In the following, we will consider a range of research objectives in the study of PEB and discuss those measurement approaches that seem most suitable to address them.

5.1 Objective 1: Characterizing individual differences in pro-environmental behavior

When approached from an individual differences perspective (Brick & Lewis, 2016; Markowitz et al., 2012), PEB measurement does not focus on quantifying properties of the behavior itself, but rather views behavior as an indicator of an individual's propensity to engage in PEB (Nelson & Hayes, 1979). When examining the personality factors that correlate with such a propensity, it may not be very promising to assess PEB in a very specific situation in the field or in the laboratory. Even if, for example, the recycling of study materials in the laboratory or of trash in the cafeteria is related to a general propensity to behave pro-environmentally, these specific instances of PEB will only reflect a small portion of the general propensity. In other words, if the propensity is measured via such specific indicators, variance in the resulting measure may primarily be error variance. Assessment of a general propensity requires a general measure of PEB and all established general measures of PEB that are currently available rely on participants' self-reports. Using the GEB scale (Kaiser, 1998) as the measure with the strongest psychometric support (Arnold, Kibbe, Hartig, & Kaiser, 2018; Kaiser & Wilson, 2004; Kaiser, Doka, Hofstetter, & Ranney, 2003; Kaiser, Merten, & Wetzel, 2018; Kaiser et al., 2001, 2010) may be a good starting point for studying the personality correlates of general PEB. However, as personality traits are typically assessed via self-reports as well, any correlations revealed by this approach are likely to be inflated by common-method variance. An obvious remedy to this problem is the use of multiple methods

to measure PEB and the supposedly related personality traits (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Adding informant reports to a study might be an easy way to achieve such methodological multiplicity (Vazire, 2006). Given the large number of established self-report measures of PEB, it is surprising that not a single established informant rating scale can be found in the PEB literature. A validated informant version of, for example, the GEB scale might substantially enhance researchers' possibilities to study individual differences in general PEB.

An intriguing alternative would be the development of laboratory test batteries or multifaceted field recordings. The latter approach is illustrated by a study by Weigel and Newman (1976) who offered participants various opportunities to engage in PEB (signing petitions, participation in a roadside litter pick-up program, recycling) over an observation period of eight months. In the laboratory, such observations of different instances of PEB from the same participants could be conducted in a more time-efficient manner (see Cornelissen et al., 2008; van Horen, van der Wal, & Grinstein, 2018). Aggregation across multiple behavioral observations into a comprehensive PEB index can help uncover relationships with personality traits or other general psychological predictors of PEB (e.g., attitude, Weigel & Newman, 1976). Along the lines of self-report scales, prospective multi-observation assessments of PEB would benefit from taking into account differences between behavioral difficulties, which might otherwise artificially reduce the correlation between behaviors (Kaiser & Wilson, 2004).

5.2 Objective 2: Understanding the mechanisms underlying pro-environmental behavior

Researchers who want to explain or change PEB are interested in causal mechanisms, and the gold standard for understanding causation is the experimental method. As discussed above, experiments on PEB are difficult to conduct in a controlled and valid way when using either self-reports or field observations to measure PEB. In contrast, behavioral experiments

in the laboratory offer the degree of experimental control that is needed to characterize causal effects on PEB. This characteristic should especially appeal to researchers who would like to develop interventions to promote PEB. Ultimately, these researchers do not wish to change PEB in the laboratory and as a consequence, it may seem intuitive to test potential interventions directly in the field. The field, however, is not the ideal situation to generate the vast amount of information that is required to understand the mechanism of action of a newly developed intervention. Due to the associated cost, field studies are often limited to the comparison of only few experimental conditions (e.g., intervention vs. control group). Such field research designs necessarily leave a large number of important questions unanswered. Does the effect size increase systematically when the intervention is administered in increasing doses? Which attributes of the intervention affect its effectiveness and which do not? How does the intervention interact with other interventions? On which situational or personal variables does its effectiveness depend? Does the intervention cause unintended side effects and which factors can help mitigate them?

Similar questions are on the mind of biomedical researchers during early (preclinical) phases of drug development. Just like researchers who want to promote PEB, they ultimately want their intervention to exert an effect in the field. However, their process of scientific inquiry naturally starts in the laboratory where they take advantage of superior experimental opportunities. Only after these opportunities have been used to establish the drug's mechanism, effectiveness, and safety in the laboratory, research would proceed to clinical phases testing the drug in the field. Such a multi-step procedure reduces the costs and increases the interpretability of late-stage field studies. Field studies do not have to examine every conceivable intervention, but they can rely on laboratory data to identify the most promising configurations of candidates and doses. If a field study conducted under limited experimental control finds an intervention to be effective, laboratory evidence for the

underlying mechanism can inform the interpretation of this effect. If a field study finds no such effect, mechanistic insights from the laboratory can inform the search for probable reasons.

Along the lines of biomedical research, research on PEB can be expected to benefit from a shift towards studying mechanisms and potential interventions under controlled laboratory conditions. In order for this research to produce meaningful results, it needs to employ validated laboratory measures of PEB. Such measures have been in short supply due to the field's focus on self-reports and field observations, but tasks such as FISH (Gifford & Gifford, 2000), the PEBT (Lange et al., 2018), or the GGG (Klein et al., 2017) are promising starting points for an experimental analysis of PEB in the laboratory.

5.3 Objective 3: Evaluating the effectiveness of interventions to promote pro-environmental behavior

While the development and evidence-based fine-tuning of PEB interventions can best be achieved under controlled laboratory conditions, interventions will ultimately have to undergo empirical evaluation in the field. In general, such evaluations should be most useful and least biased when they involve the unobtrusive observation of PEB or of a strongly correlated PEB product. Field evaluations may address different questions about a given intervention and these questions favor different observation methods. For example, one could ask whether an intervention that proved effective in the laboratory also promotes PEB when participants do not know that they participate in a study. In this case, it might be advisable to contrive a situation (e.g., distribute flyers and track how participants dispose of them) in order to increase the power of the study. Alternatively, researchers may be interested in examining whether the effects of a particular intervention are strong enough to stand out from the noise typical for a particular PEB. This question cannot be addressed in a contrived situation that involves the reduction of behavior-typical noise and rather calls for the observation of PEB in

naturalistic conditions. Both these kinds of research would benefit from the availability of established field assessment protocols. By this means, different interventions can be directly compared based on the effects they exert on PEB observed with the same coding system in a similar situation.

When using self-reports to evaluate the effectiveness of interventions in the field, researchers may obtain more accurate and sensitive data when employing diary procedures or ecological momentary assessment (EMA, Shiffman, Stone, & Hufford, 2008) instead of global retrospective report formats. In EMA, participants can report their behavior in a direct response to a prompt that they receive on a mobile device. In comparison to conventional self-report scales, these assessment procedures typically refer to a smaller time frame and a more narrowly circumscribed behavior. Respondents might thus have less difficulty in accurately recalling behavioral properties (Shiffman et al., 2008). By allowing aggregation across multiple occasions, such procedures might also contribute to the reduction of measurement error (Epstein, 1979). When validated, such methods may complement behavioral observations in evaluating the effects of PEB interventions. In contrast, global retrospective self-report scales might be most suitable when an intervention is designed to change an underlying propensity to engage in PEB (rather than a particular PEB itself).

6. Conclusion

Researchers interested in measuring PEB can choose from a large number of assessment approaches and, in all likelihood, this number will continue to grow in the future. Different measures lend themselves to different kinds of research questions. Global aggregate measures might be best-suited for research on individual differences in PEB, whereas the mechanisms underlying PEB can best be elucidated using laboratory tasks. Laboratory assessment also allows for the development and fine-tuning of interventions, the effectiveness of which can ultimately be assessed by means of field observations of PEB. Independent of

the measurement approach, PEB researchers should strive to develop and select assessment tools based on evaluations of their psychometric properties. By using and building on established and validated measures, researchers contribute to a cumulative research culture that will improve our understanding of PEB in the long run.

References

Alisat, S., & Riemer, M. (2015). The environmental action scale: Development and psychometric evaluation. *Journal of Environmental Psychology*, *43*, 13-23. https://doi.org/10.1016/j.jenvp.2015.05.006

Alpizar, F., Carlsson, F., & Johansson-Stenman, O. (2008). Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. *Journal of Public Economics*, *92*, 1047-1060. https://doi.org/10.1016/j.jpubeco.2007.11.004

Arnold, O., Kibbe, A., Hartig, T., & Kaiser, F.G. (2018). Capturing the environmental impact of individual lifestyles: evidence of the criterion validity of the general ecological behavior scale. Environment and Behavior, *50*, 350-372. https://doi.org/10.1177/0013916517701796

Bamberg, S. (2002). Effects of implementation intentions on the actual performance of new environmentally friendly behaviours—results of two field experiments. *Journal of Environmental Psychology*, *22*, 399-411. https://doi.org/10.1006/jevp.2002.0278

Bamberg, S. (2006). Is a residential relocation a good opportunity to change people's travel behavior? Results from a theory-driven intervention study. *Environment and Behavior*, *38*, 820-840. https://doi.org/10.1177/0013916505285091

Barber, N.A., Bishop, M., & Gruen, T. (2014). Who pays more (or less) for pro-environmental consumer goods? Using the auction method to assess actual willingness-to-pay. *Journal of Environmental Psychology*, *40*, 218-227. https://doi.org/10.1016/j.jenvp.2014.06.010

Bergquist, M., & Nilsson, A. (2016). I saw the sign: promoting energy conservation via normative prompts. *Journal of Environmental Psychology*, *46*, 23-31. https://doi.org/10.1016/j.jenvp.2016.03.005

Bleys, B., Defloor, B., Van Ootegem, L., & Verhofstadt, E. (2018). The environmental impact of individual behavior: self-assessment versus the ecological footprint. *Environment and Behavior, 50,* 187-212. https://doi.org/10.1177/0013916517693046

Boiral, O., & Paillé, P. (2012). Organizational citizenship behaviour for the environment: Measurement and validation. *Journal of Business Ethics*, *109*, 431-445. http://dx.doi.org/10.1007/s10551-011-1138-9

Bolderdijk, J.W., Knockaert, J., Steg, E.M., & Verhoef, E.T. (2011). Effects of Pay-As-You-Drive vehicle insurance on young drivers' speed choice: Results of a Dutch field experiment. *Accident Analysis & Prevention*, *43*, 1181-1186. https://doi.org/10.1016/j.aap.2010.12.032

Brick, C., & Lewis, G.J. (2016). Unearthing the "green" personality: Core traits predict environmentally friendly behavior. *Environment and Behavior, 48,* 635-658. https://doi.org/10.1177/0013916514554695

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*, 513-531. http://dx.doi.org/10.1037/0003-066X.32.7.513

Brown, B.B., Werner, C.M., Smith, K.R., Tribby, C.P., Miller, H.J., Jensen, W.A., & Tharp, D. (2016). Environmental, behavioral, and psychological predictors of transit ridership: Evidence from a community intervention. *Journal of Environmental Psychology*, *46*, 188-196. https://doi.org/10.1016/j.jenvp.2016.04.010

Chao, Y.L., & Lam, S.P. (2011). Measuring responsible environmental behavior: Self-reported and other-reported measures and their differences in testing a behavioral model. *Environment and Behavior*, *43*, 53-71. https://doi.org/10.1177/0013916509350849

649 Chu, P.Y., & Chiu, J.F. (2003). Factors influencing household waste recycling behavior: test
650 of an integrated model. *Journal of Applied Social Psychology*, *33*, 604-626.
651 https://doi.org/10.1111/j.1559-1816.2003.tb01915.x
652 Churchill Jr, G.A. (1979). A paradigm for developing better measures of marketing
653 constructs. *Journal of Marketing Research*, *16*, 64-73.
654 https://doi.org/10.2307/3150876
655 Clements, J.M., McCright, A.M., Dietz, T., & Marquart-Pyatt, S.T. (2015). A behavioural
656 measure of environmental decision-making for social surveys. *Environmental*
657 *Sociology*, *1*, 27-37. https://doi.org/10.1080/23251042.2015.1020466
658 Cornelissen, G., Pandelaere, M., Warlop, L., & Dewitte, S. (2008). Positive cueing:
659 Promoting sustainable consumer behavior by cueing common environmental
660 behaviors as environmental. *International Journal of Research in Marketing*, *25*, 46-
661 55. https://doi.org/10.1016/j.ijresmar.2007.06.002
662 Corral-Verdugo, V., Bernache, G., Encinas, L., & Garibaldi, L.C. (1994-95). A comparison of
663 two measures of reuse and recycling behavior: Self-report and material culture.
664 *Journal of Environmental Systems*, *23*, 313-313. https://doi.org/10.2190/EJP5-56TU-
665 0G9U-FUJL
666 Cote Jr, J.A. (1984). Use of household refuse analysis to measure usual and period-specific
667 food consumption. *American Behavioral Scientist*, *28*, 129-138.
668 https://doi.org/10.1177/000276484028001010
669 Dono, J., Webb, J., & Richardson, B. (2010). The relationship between environmental
670 activism, pro-environmental behaviour and social identity. *Journal of Environmental*
671 *Psychology*, *30*, 178-186. https://doi.org/10.1016/j.jenvp.2009.11.006
672 Dunlosky, J., Bottiroli, S., & Hartwig, M. (2009). Sins committed in the name of ecological
673 validity: A call for representative design in education science. In D.J. Hacker, J.

Dunlosky, & A.C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 430–440). New York, NY: Routledge.

Dupré, M., & Meineri, S. (2016). Increasing recycling through displaying feedback and social comparative feedback. *Journal of Environmental Psychology*, *48*, 101-107. https://doi.org/10.1016/j.jenvp.2016.07.004

Dwyer, P.C., Maki, A., & Rothman, A.J. (2015). Promoting energy conservation behavior in public settings: The influence of social norms and personal responsibility. *Journal of Environmental Psychology*, *41*, 30-34. https://doi.org/10.1016/j.jenvp.2014.11.002

Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology, 37*, 1097-1126. http://dx.doi.org/10.1037/0022-3514.37.7.1097

Evans, G.W., Brauchle, G., Haq, A., Stecker, R., Wong, K., & Shapiro, E. (2007). Young children's environmental attitudes and behaviors. *Environment and Behavior, 39*, 635-658. https://doi.org/10.1177/0013916506294252

Foxx, R.M., & Hake, D.F. (1977). Gasoline conservation: A procedure for measuring and reducing the driving of college students. *Journal of Applied Behavior Analysis, 10*, 61-74. https://doi.org/10.1901/jaba.1977.10-61

Gamba, R. J., & Oskamp, S. (1994). Factors influencing community residents' participation in commingled curbside recycling programs. *Environment and Behavior, 26*, 587-612. https://doi.org/10.1177/0013916594265001

Gatersleben, B., Steg, L., & Vlek, C. (2002). Measurement and determinants of environmentally significant consumer behavior. *Environment and Behavior, 34*, 335-362. https://doi.org/10.1177/0013916502034003004

Geller, E.S., Farris, J.C., & Post, D.S. (1973). Prompting a consumer behavior for pollution control. *Journal of Applied Behavior Analysis, 6*, 367-376. https://doi.org/10.1901/jaba.1973.6-367

Geller, E.S., Witmer, J.F., & Tuso, M.A. (1977). Environmental interventions for litter control. *Journal of Applied Psychology, 62*, 344-351. http://dx.doi.org/10.1037/0021-9010.62.3.344

Gifford, J., & Gifford, R. (2000). FISH 3: A microworld for studying social dilemmas and resource management. *Behavior Research Methods, Instruments, & Computers, 32*, 417-422. https://doi.org/10.3758/BF03200810

Gifford, R. (2014). Environmental psychology matters. *Annual Review of Psychology, 65*, 541-579. https://doi.org/10.1146/annurev-psych-010213-115048

Gifford, R., & Hine, D.W. (1997). Toward cooperation in commons dilemmas. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement, 29*, 167-179. http://dx.doi.org/10.1037/0008-400X.29.3.167

Gifford, R., & Wells, J. (1991). FISH: A commons dilemma simulation. *Behavior Research Methods, Instruments, & Computers, 23*, 437-441. https://doi.org/10.3758/BF03203408

Goldstein, N.J., Cialdini, R.B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research, 35*, 472-482. https://www.jstor.org/stable/10.1086/586910

Gosling, S. ., Ko, S. ., Mannarelli, T., & Morris, M. . (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82*, 379-398. http://dx.doi.org/10.1037/0022-3514.82.3.379

Gregory, G.D., & Di Leo, M. (2003). Repeated behavior and environmental psychology: the role of personal involvement and habit formation in explaining water consumption.

Journal of Applied Social Psychology, *33*, 1261-1296. https://doi.org/10.1111/j.1559-1816.2003.tb01949.x

Hake, D.F., & Zane, T. (1981). A community-based gasoline conservation project: Practical and methodological considerations. *Behavior Modification*, *5*, 435-458. https://doi.org/10.1177/014544558154001

Hamann, K.R., Reese, G., Seewald, D., & Loeschinger, D.C. (2015). Affixing the theory of normative conduct (to your mailbox): Injunctive and descriptive norms as predictors of anti-ads sticker use. *Journal of Environmental Psychology*, *44*, 1-9. https://doi.org/10.1016/j.jenvp.2015.08.003

Hanss, D., & Böhm, G. (2013). Promoting purchases of sustainable groceries: An intervention study. *Journal of Environmental Psychology*, *33*, 53-67. https://doi.org/10.1016/j.jenvp.2012.10.002

Huddart Kennedy, E., Krahn, H., & Krogman, N.T. (2015). Are we counting what counts? A closer look at environmental concern, pro-environmental behaviour, and carbon footprint. *Local Environment*, *20*, 220-236. https://doi.org/10.1080/13549839.2013.837039

Huffman, A.H., Van Der Werff, B.R., Henning, J.B., & Watrous-Rodriguez, K. (2014). When do recycling attitudes predict recycling? An investigation of self-reported versus observed behavior. *Journal of Environmental Psychology*, *38*, 262-270. https://doi.org/10.1016/j.jenvp.2014.03.006

Jacobs, M.H., & Harms, M. (2014). Influence of interpretation on conservation intentions of whale tourists. *Tourism Management*, *42*, 123-131. https://doi.org/10.1016/j.tourman.2013.11.009

Kaiser, F.G. (1998). A general measure of ecological behavior. *Journal of Applied Social Psychology*, *28*, 395-422. https://doi.org/10.1111/j.1559-1816.1998.tb01712.x

Kaiser, F.G., Byrka, K., & Hartig, T. (2010). Reviving Campbell's paradigm for attitude research. *Personality and Social Psychology Review, 14*, 351-367. https://doi.org/10.1177/1088868310366452

Kaiser, F.G., Doka, G., Hofstetter, P., & Ranney, M.A. (2003). Ecological behavior and its environmental consequences: A life cycle assessment of a self-report measure. *Journal of Environmental Psychology*, *23*, 11-20. https://doi.org/10.1016/S0272-4944(02)00075-0

Kaiser, F. G., Frick, J., & Stoll-Kleemann, S. (2001). Zur Angemessenheit selbstberichteten Verhaltens: Eine Validitätsuntersuchung der Skala Allgemeinen Ökologischen Verhaltens [Accuracy of self-reports: Validating the general ecological behavior scale]. *Diagnostica*, *47*, 88–95. https://doi.org/10.1026//0012-1924.47.2.88

Kaiser, F. G., Merten, M., & Wetzel, E. (2018). How do we know we are measuring environmental attitude? Specific objectivity as the formal validation criterion for measures of latent attributes. *Journal of Environmental Psychology*, *55*, 139-146. https://doi.org/10.1016/j.jenvp.2018.01.003

Kaiser, F.G., Oerke, B., & Bogner, F.X. (2007). Behavior-based environmental attitude: Development of an instrument for adolescents. *Journal of Environmental Psychology*, 27, 242-251. https://doi.org/10.1016/j.jenvp.2007.06.004

Kaiser, F.G., & Wilson, M. (2004). Goal-directed conservation behavior: The specific composition of a general performance. *Personality and Individual Differences*, *36*, 1531-1544. https://doi.org/10.1016/j.paid.2003.06.003

Katzev, R., & Bachman, W. (1982). Effects of deferred payment and fare manipulations on urban bus ridership. *Journal of Applied Psychology, 67*, 83-88. http://dx.doi.org/10.1037/0021-9010.67.1.83

Kazdin, A.E. (1979). Unobtrusive measures in behavioral assessment. *Journal of Applied Behavior Analysis*, *12*, 713-724. http://dx.doi.org/10.1901/jaba.1979.12-713

Kazdin, A.E. (1982). Observer effects: Reactivity of direct observation. *New Directions for Methodology of Social & Behavioral Science*, *14*, 5-19.

Keller, J. J. (1991). The recycling solution: How I increased recycling on Dilworth Road. *Journal of Applied Behavior Analysis*, *24*, 617-619. https://doi.org/10.1901/jaba.1991.24-617

Klein, S.A., & Hilbig, B.E. (2018). How virtual nature experiences can promote pro-environmental behavior. *Journal of Environmental Psychology*, *60*, 41-47. https://doi.org/10.1016/j.jenvp.2018.10.001

Klein, S. A., & Hilbig, B. E. (2019). On the lack of real consequences in consumer choice research: And its consequences. *Experimental Psychology*, *66*, 68-76. https://doi.org/10.1027/1618-3169/a000420

Klein, S.A., Hilbig, B.E., & Heck, D.W. (2017). Which is the greater good? A social dilemma paradigm disentangling environmentalism and cooperation. *Journal of Environmental Psychology*, *53*, 40-49. https://doi.org/10.1016/j.jenvp.2017.06.001

Kollmuss, A., & Agyeman, J. (2002). Mind the gap: why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental Education Research*, *8*, 239-260. https://doi.org/10.1080/13504620220145401

Kormos, C., & Gifford, R. (2014). The validity of self-report measures of proenvironmental behavior: A meta-analytic review. *Journal of Environmental Psychology*, *40*, 359-371. https://doi.org/10.1016/j.jenvp.2014.09.003

Lam, S.P., & Cheng, S.I. (2002). Cross-informant agreement in reports of environmental behavior and the effect of cross-questioning on report accuracy. *Environment and Behavior*, *34*, 508-520. https://doi.org/10.1177/00116502034004006

Lange, F., Steinke, A., & Dewitte, S. (2018). The Pro-Environmental Behavior Task: A laboratory measure of actual pro-environmental behavior. *Journal of Environmental Psychology*, *56*, 46-54. https://doi.org/10.1016/j.jenvp.2018.02.007

Larson, L.R., Stedman, R.C., Cooper, C.B., & Decker, D.J. (2015). Understanding the multi-dimensional structure of pro-environmental behavior. *Journal of Environmental Psychology*, *43*, 112-124. https://doi.org/10.1016/j.jenvp.2015.06.004

Lehman, P. K., & Geller, E. S. (2004). Behavior analysis and environmental protection: Accomplishments and potential for more. *Behavior and Social Issues*, *13*, 13-32. https://doi.org/10.5210/bsi.v13i1.33

Maki, A., & Rothman, A.J. (2017). Understanding proenvironmental intentions and behaviors: The importance of considering both the behavior setting and the type of behavior. *The Journal of Social Psychology*, *157*, 517-531. https://doi.org/10.1080/00224545.2016.1215968

Maleetipwan-Mattsson, P., Laike, T., & Johansson, M. (2013). Self-report diary: A method to measure use of office lighting. *Leukos*, *9*, 291-306. https://doi.org/10.1582/LEUKOS.2013.09.04.004

Mallett, R.K., & Melchiori, K.J. (2016). Creating a water-saver self-identity reduces water use in residence halls. *Journal of Environmental Psychology*, *47*, 223-229. https://doi.org/10.1016/j.jenvp.2016.07.001

Markle, G.L. (2013). Pro-environmental behavior: Does it matter how it's measured? Development and validation of the Pro-Environmental Behavior Scale (PEBS). *Human Ecology*, *41*, 905-914. https://doi.org/10.1007/s10745-013-9614-8

Markowitz, E.M., Goldberg, L.R., Ashton, M.C., & Lee, K. (2012). Profiling the "pro-environmental individual": A personality perspective. *Journal of Personality*, *80*, 81-111. https://doi.org/10.1111/j.1467-6494.2011.00721.x

821 Mayer, J., & Geller, E.S. (1982-1983). Motivating energy efficient travel: A community-
822 based intervention for encouraging biking. *Journal of Environmental Systems*, *12*, 99-
823 111. https://doi.org/10.2190/C9H7-6ULX-W52K-HW2D
824 McCall, G. J. (1984). Systematic field observation. *Annual Review of Sociology*, *10*, 263-282.
825 http://dx.doi.org/10.1146/annurev.so.10.080184.001403
826 Meleady, R., Abrams, D., Van de Vyver, J., Hopthrow, T., Mahmood, L., Player, A., ... &
827 Leite, A.C. (2017). Surveillance or self-surveillance? Behavioral cues can increase the
828 rate of drivers' pro-environmental behavior at a long wait stop. *Environment and*
829 *Behavior*, *49*(, 1156-1172. https://doi.org/10.1177/0013916517691324
830 Milfont, T. L. (2009). The effects of social desirability on self-reported environmental
831 attitudes and ecological behaviour. *The Environmentalist*, *29*, 263-269.
832 https://doi.org/10.1007/s10669-008-9192-2
833 Murtagh, N., Gatersleben, B., Cowen, L., & Uzzell, D. (2015). Does perception of automation
834 undermine pro-environmental behaviour? Findings from three everyday settings.
835 *Journal of Environmental Psychology*, *42*, 139-148.
836 https://doi.org/10.1016/j.jenvp.2015.04.002
837 Nelson, R. O., & Hayes, S. C. (1979). The nature of behavioral assessment: A commentary.
838 *Journal of Applied Behavior Analysis*, *12*, 491-500.
839 https://doi.org/10.1901/jaba.1979.12-491
840 Norton, T.A., Parker, S.L., Zacher, H., & Ashkanasy, N.M. (2015). Employee green behavior:
841 A theoretical framework, multilevel review, and future research agenda. *Organization*
842 *& Environment*, *28*, 103-125. https://doi.org/10.1177/1086026615575773
843 Obery, A., & Bangert, A. (2017). Exploring the influence of nature relatedness and perceived
844 science knowledge on proenvironmental behavior. *Education Sciences*, *7*: 17.
845 https://doi.org/10.3390/educsci7010017

Peattie, K. (2010). Green consumption: behavior and norms. *Annual Review of Environment and Resources*, *35*, 195-228. https://doi.org/10.1146/annurev-environ-032609-094328

Pisano, I., & Lubell, M. (2017). Environmental behavior in cross-national perspective: A multilevel analysis of 30 countries. *Environment and Behavior*, *49*, 31-58. https://doi.org/10.1177/0013916515600494

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., & Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879-903. http://dx.doi.org/10.1037/0021-9010.88.5.879

Roberts, J.A. (1996). Green consumers in the 1990s: profile and implications for advertising. *Journal of Business Research*, *36*, 217-231. https://doi.org/10.1016/0148-2963(95)00150-6

Robertson, J.L., & Barling, J. (2017). Toward a new measure of organizational environmental citizenship behavior. *Journal of Business Research*, *75*, 57-66. https://doi.org/10.1016/j.jbusres.2017.02.007

Sapci, O., & Considine, T. (2014). The link between environmental attitudes and energy consumption behavior. *Journal of Behavioral and Experimental Economics*, *52*, 29-34. https://doi.org/10.1016/j.socec.2014.06.001

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. Infancy, *2*, 419-436. https://doi.org/10.1207/S15327078IN0204_02

Schultz, P.W., Messina, A., Tronu, G., Limas, E.F., Gupta, R., & Estrada, M. (2016). Personalized normative feedback and the moderating role of personal norms: A field experiment to reduce residential water consumption. *Environment and Behavior*, *48*, 686-710. https://doi.org/10.1177/0013916514553835

Seebauer, S., Fleiß, J., & Schweighart, M. (2017). A household is not a person: Consistency of pro-environmental behavior in adult couples and the accuracy of proxy-reports. *Environment and Behavior*, *49*, 603-637. https://doi.org/10.1177/0013916516663796

Shiffman, S., Stone, A.A., & Hufford, M.R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1-32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Siero, F.W., Bakker, A.B., Dekker, G.B., & Van Den Burg, M.T. (1996). Changing organizational energy consumption behaviour through comparative feedback. *Journal of Environmental Psychology*, *16*, 235-246. https://doi.org/10.1006/jevp.1996.0019

Simmons, J.P., Nelson, L.D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359-1366. https://doi.org/10.1177/0956797611417632

Smith, G.A. (1986). Observer drift: A drifting definition. *The Behavior Analyst*, *9*, 127-128.

Sörqvist, P., Haga, A., Holmgren, M., & Hansla, A. (2015). An eco-label effect in the built environment: Performance and comfort effects of labeling a light source environmentally friendly. *Journal of Environmental Psychology*, *42*, 123-127. https://doi.org/10.1016/j.jenvp.2015.03.004

Steg, L., & Vlek, C. (2009). Encouraging pro-environmental behaviour: An integrative review and research agenda. *Journal of Environmental Psychology*, *29*, 309-317. https://doi.org/10.1016/j.jenvp.2008.10.004

Stern, P.C. (1992). Psychological dimensions of global environmental change. *Annual Review of Psychology*, *43*, 269-302. https://doi.org/10.1146/annurev.ps.43.020192.001413

Stern, P. C. (2000). New environmental theories: toward a coherent theory of environmentally significant behavior. *Journal of Social Issues*, *56*, 407-424. https://doi.org/10.1111/0022-4537.00175

Stern, P. C., Dietz, T., Abel, T., Guagnano, G. A., & Kalof, L. (1999). A value-belief-norm theory of support for social movements: The case of environmentalism. *Human Ecology Review*, *6*, 81-97.

Swim, J.K., Clayton, S., & Howard, G.S. (2011). Human behavioral contributions to climate change: Psychological and contextual drivers. *American Psychologist*, *66*, 251-264. http://dx.doi.org/10.1037/a0023472

Tam, K. P., & Chan, H. W. (2017). Environmental concern has a weaker association with pro-environmental behavior in some societies than others: A cross-cultural psychology perspective. *Journal of Environmental Psychology*, *53*, 213-223. https://doi.org/10.1016/j.jenvp.2017.09.001

Tarditi, C., Hahnel, U.J., Jeanmonod, N., Sander, D., & Brosch, T. (in press). Affective Dilemmas: The Impact of Trait Affect and State Emotion on Sustainable Consumption Decisions in a Social Dilemma Task. *Environment and Behavior*. https://doi.org/10.1177/0013916518787590

Taube, O., Kibbe, A., Vetter, M., Adler, M., & Kaiser, F.G. (2018). Applying the Campbell Paradigm to sustainable travel behavior: Compensatory effects of environmental attitude and the transportation environment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *56*, 392-407. https://doi.org/10.1016/j.trf.2018.05.006

Truelove, H.B., & Gillis, A.J. (2018). Perception of pro-environmental behavior. *Global Environmental Change*, *49*, 175-185. https://doi.org/10.1016/j.gloenvcha.2018.02.009

van Horen, F., van der Wal, A., & Grinstein, A. (2018). Green, Greener, Greenest: Can Competition Increase Sustainable Behavior? *Journal of Environmental Psychology*, *59*, 16-25. https://doi.org/10.1016/j.jenvp.2018.08.007

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, *40*, 472-481. https://doi.org/10.1016/j.jrp.2005.03.003

Vesely, S., & Klöckner, C.A. (2018). How anonymity and norms influence costly support for environmental causes. *Journal of Environmental Psychology*, *58*, 27-30. https://doi.org/10.1016/j.jenvp.2018.07.013

Vining, J., Ebreo, A., Bechtel, R.B., & Churchman, A. (2002). Emerging theoretical and methodological perspectives on conservation behaviour. In R.B. Bechtel & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 541-558). Hoboken, NJ: Wiley.

Walker, J.M. (1979). Energy demand behavior in a master-metered apartment complex: An experimental analysis. *Journal of Applied Psychology*, *64*, 190-196. http://dx.doi.org/10.1037/0021-9010.64.2.190

Weigel, R.H., & Newman, L.S. (1976). Increasing attitude-behavior correspondence by broadening the scope of the behavioral measure. *Journal of Personality and Social Psychology*, *33*, 793-802. http://dx.doi.org/10.1037/0022-3514.33.6.793

Whitmarsh, L. (2009). Behavioural responses to climate change: Asymmetry of intentions and impacts. *Journal of Environmental Psychology*, *29*, 13-23. https://doi.org/10.1016/j.jenvp.2008.05.003

Wilson E.O. (1988). The current state of biological diversity. In E.O. Wilson & F.M. Peter (Eds.), *Biodiversity* (pp. 3-18), Washington, DC: National Academy Press.

Winett, R.A., & Nietzel, M.T. (1975). Behavioral ecology: Contingency management of consumer energy use. *American Journal of Community Psychology*, *3*, 123-133. https://doi.org/10.1007/BF00877787

Wynes, S., & Nicholas, K.A. (2017). The climate mitigation gap: education and government recommendations miss the most effective individual actions. *Environmental Research Letters*, *12*: 074024. https://doi.org/10.1088/1748-9326/aa7541

Yoder, P., Lloyd, B. P., & Symons, F. (2018). *Observational measurement of behavior*. Baltimore, MD: Paul H. Brookes Publishing.

Xiao, G., Juan, Z., & Zhang, C. (2015). Travel mode detection based on GPS track data and Bayesian networks. *Computers, Environment and Urban Systems*, *54*, 14-22. https://doi.org/10.1016/j.compenvurbsys.2015.05.005

Supplementary Materials

Established self-report scales for the assessment of pro-environmental behavior

In the following, we present a literature review of 33 established multi-item scales for the assessment of pro-environmental behavior (Table S1). These scales were identified by searching the reference lists of articles presenting newly developed pro-environmental behavior scales and by screening the Google Scholar records that cited these articles. While this approach likely resulted in the identification of most scales that are embedded in the respective research literature, we cannot guarantee the completeness of the list provided in Table S1. Nonetheless, we think that this list can be helpful for researchers looking for a measure to use in their studies.

Measures were included as being "established" when they had undergone psychometric evaluation in their development or when they had been used in multiple studies in the exact same form. Please note that this does neither imply that all of the scales listed below have been adequately evaluated nor that all psychometric evaluations yielded favorable results.

Of the identified scales, 20 were designed to measure pro-environmental behavior in general, while the remaining 13 focus on more or less specific domains of pro-environmental behavior. Domain-general measures of pro-environmental behavior can further be differentiated based on their dimensionality. Unidimensional and multidimensional conceptualizations of pro-environmental behavior were found in ten studies each. This division is reflective of the ongoing debate about the similarities and differences between pro-environmental behaviors (Kaiser, 1998; Larson et al., 2015; Lee, Kim, Kim, & Choi, 2014; Stern, 2000; Vining & Ebreo, 2002). Whether pro-environmental behavior is multidimensional or not obviously depends on one's operational definition of unidimensionality. Consider, for example, the General Ecological Behavior (GEB) scales developed by Kaiser and colleagues which are portrayed as being unidimensional in Table S1. In fact, evaluations of the factor structure of these scales have indicated that a six-dimensional model fits the data significantly better than a unidimensional one (Kaiser, Oerke, & Bogner, 2007; Kaiser & Wilson, 2004). These results indicate that there are systematic differences between, for example, pro-environmental behavior in the domain of energy conservation and pro-environmental behavior in the domain of recycling. However, despite these differences, there were also strong correlations between

the six dimensions of pro-environmental behavior, and the amount of information that was lost by adopting a unidimensional model was judged to be negligible (Kaiser et al., 2007; Kaiser & Wilson, 2004). Hence, different pro-environmental behaviors appear to be related enough to be combined into a meaningful domain-general index while also being different enough to create domain-specific subscales.

Of note, the large number of different domain-general self-report scales suggests that many researchers prefer to create their own measure of pro-environmental behavior over using already existing scales. In many cases, the incremental value of newly developed measures and their relationship to established measures remain unspecified. In other cases, the development of new measures seems clearly warranted, for example, when researchers wish to assess pro-environmental behavior in specific populations such as students in secondary education (Kaiser et al., 2007) or even younger children (Evans et al., 2007).

Similar to these population-specific measures, domain-specific measures of pro-environmental behavior might be important additions to the toolbox of the pro-environmental behavior researcher. Some of these measures focus on pro-environmental behavior that is shown in a particular context (e.g., the workplace Boiral & Paillé, 2012; Robertson & Barling, 2017). Other scales are designed to capture a particular facet of pro-environmental behavior (e.g., activism, Alisat & Riemer, 2015). These facets can be rather broad (e.g., consumption behavior, Roberts, 1996) or very specific (e.g., littering, Ojedokun, 2016).

Table S1
Overview of established self-report scales for the measurement of pro-environmental behavior

	domain	development context	items	α_{tot}	dim	subscale	example item	α_{sub}	correlates
domain-general measures									
Recurring Pro-Environmental Behavior Scale (Brick, Sherman, & Kim, 2017)[a]	general	MTurkers, USA	21	.82-.87	1		How often do you turn your personal electronics off or in low-power mode when not in use?		environmentalist identity, environmental attitudes, climate change beliefs, attitudes about environmentalists
Ecological Behaviour Scale (Casey & Scott, 2006)[a]	general	Students, Australia	17	.81	1		I use the washing machine only when it has a full load.		environmental concern
Children's Environmental Behavior Jumping Game (Evans et al., 2007)[b]	general	pupils, 7 yrs, USA	8	.49*	1		How often do you leave the refrigerator door open while deciding what to eat?		mother-rated PEB, NOT: environmental attitudes
Environmentally Responsible Behavior Scale (Iwata, 2001)	general	students, Japan	15	.71	1		I leave my TV set turned on while I am busy elsewhere.		environmental attitudes
General Ecological Behavior Scale (Kaiser, 1998; Kaiser & Wilson, 2004)[c]	general	members of transportation associations, Switzerland, repeated in several countries	30-65	.72-.88 .71-88*	1		I wash dirty clothes without prewashing.		willingness to behave pro-environmentally, acceptance of governmental prohibitions, membership in environmental organization, actual PEB
General Ecological Behavior Scale – adolescent version (Kaiser, Oerke, & Bogner, 2007)[b]	general	pupils 9-18 yrs, Germany	40	.78 .80*	1		I insist on holidays close to home.		environmental attitudes, other self-reported PEB measure
Actual Commitment to Ecological Behavior (Maloney & Ward, 1973)[a,b]	general	members of environmental organization,	36	.92	1		I keep track of my congressman and senator's voting records on environment issues.		membership in environmental organization, attitude, verbal

		students, non-students, USA							commitment, NOT: environmental knowledge
Actual Commitment to Ecological Behavior – revised (Maloney, Ward, & Braucht, 1975)[a]	general	members of environmental organization, students, non-students, USA	10	.89	1		I keep track of my congressman and senator's voting records on environment issues.		membership in environmental organization, attitude, verbal commitment, NOT: environmental knowledge
Environmental Behavior Scale (Schultz et a., 2005)	general	students, Brazil, the Czech Republic, Germany, India, New Zealand, Russia	10	.60-.75	1		How often have you looked for ways to reuse things in the past year?		biospheric values
ECOSCALE- Action Taken (Stone, Barnes, Montgomery, 1995)	general	students, USA	5	n/p	1		I turn in polluters when I see them dumping toxic liquids.		other self-reported PEB measure
Pro-Environmental Activities Scale (Tilikidou, Adamson, & Sarmaniotis, 2002)	general	residents, Greece	11	.76-.80	2				other self-reported PEB measure, environmental attitudes
						participative activities	I often take part into environmental protection events.	.79-.87	
						individual activities	I try to use less water.	.67-.70	
Environmentalism Scale (Stern, Diez, Abel, Guagnano, & Kalof, 1999)	general	residents, USA	14	n/p	3				personal environmental norm

						consumer behavior	How often do you avoid buying products from a company that you know may be harming the environment?	.72	
						willingness to sacrifice	I would be willing to pay much higher taxes in order to protect the environment.	.78	
						environmental citizenship	In the last twelve months, have you read any newsletters, magazines or other publications written by environmental groups?	.77	
Environmental Behavior Scale (Karp, 1996)[b]	general	students, USA	15	.82	3				biospheric values
						good citizen	I tried not to litter.	.72	
						activist	I contributed money to an environmental group.	.70	
						heathy consumer	I bought organically grown produce.	.69	
Pro-Environmental Behavior Scale (Larson, Stedman, Cooper, & Decker, 2015)[b]	general	residents, USA	13	n/p	4				
						conservation lifestyle	I conserved water or energy in my home.	.79	
						land stewardship	I made my yard or my land more desirable for wildlife.	.64	
						social environmentalism	I talked to others in my community about environmental issues.	.78	
						environmental citizenship	I signed a petition about an environmental issue.	.84	
Pro-Environmental Behavior Scale (Markle, 2013)	general	students, USA	19	.76-.80	4				environmental concern, environmental identity, support for environmental regulation, other self-reported PEB measure
						conservation	How often do you limit your time in the shower in order to conserve water?	.74-.77	
						environmental citizenship	Are you currently a member of any environmental, conservation, or wildlife protection group?	.63-.65	

						food	During the past year have you decreased the amount of beef you consume?	.66-.67	
						transportation	During the past year how often have you car-pooled?	.62-.64	
Children's Responsible Environmental Behavior Scale (Erdogan, Ok, & Marcinkowski, 2012)	general	pupils, 10-11 yrs, Turkey	23	n.p.	4				
						political action	I talked to government officials in order to enforce environmental laws or punish people who violate these laws.	.92	
						physical action/eco-management	I picked up litter, trash, and garbage in schools, picnic areas, parks, and street and threw them in garbage bins.	.70	
						consumer and economic action	I purchased products which are recyclable and which are made from recycled materials (e.g. I purchased the products on which there is a recycling sign).	.70	
						individual and public persuasion	I talked with my friends about what measures to be taken to protect and not harm the environment.	.80	
Behavior Inventory of Environmental Action (Sia, Hunderford, & Tomera, 1986)[a,b]	general	members of environmental organization and educational travelers. USA	n/p	.90	5				environmental skills and knowledge, attitude towards pollution, environmental sensitivity
						eco-management	n/p	n/p	
						persuasion	n/p	n/p	
						consumerism	n/p	n/p	
						political action	n/p	n/p	
						legal action	n/p	n/p	
Environmental Behavior Scale (Stanley, Lasonde, & Weiss, 1996)	general	students, USA	37	.90	6				environmental concern
						purchase	How often do you buy biodegradable laundry soap?	n/p	
						recycling	How often do you recycle plastics	n/p	

						maintenance	How often do you keep tires inflated?	n/p	
						curtailment	How often do you turn air conditioning down?	n/p	
						transportation efficiency	How often do you avoid driving?	n/p	
							How often do you install toilet dams?	n/p	
Tourists' Environmentally Responsible Behavior (Lee, Jan, & Yang, 2013)	general (and tourism-specific)	tourists, Taiwan	24	n/p	7				
						civil action	I join in community cleanup efforts	.84	
						financial action	I buy environmentally friendly products	.82	
						physical action	I turn off lights if I am leaving a room for more than 10 min	.79	
						persuasive action	I convince someone to buy fruits and vegetables loose rather than in plastic bags	.87	
						sustainable behavior	I observe the history and culture heritage detailed.	.83	
						pro-environmental behavior	I voluntarily stop visiting a favorite spot if it needed to recover from environmental damage.	.81	
						environmentally friendly behavior	I tell my companions not to feed the animals.	.77	
Stanford Climate Change Behavior Survey (Armel, Yan, Todd, & Robinson, 2011) [a]	general, greenhouse-gas-relevant behaviors	students, USA	97	n/p	10				importance of environmental sustainability, membership in environmental organizations, being vegetarian, NOT: number of environmental classes
						electricity	How many times per week do you usually shower?	.66	
						high GHG transport	Approximately how often do you check your car tire inflation pressure?	.64	
						low GHG transport	Mark the answer corresponding to the number of one way trips per week you typically travel by bus.	n/p	

						no GHG transport	Mark the answer corresponding to the number of one way trips per week you typically travel by biking.	n/p	
						waste	How much trash do you personally produce each week in the place where you live? Estimate how many plastic grocery bags of trash you would fill.	.81	
						food packaging	How often on average have you eaten 3–4 oz. canned fish during the past month?	.51	
						high GHG foods	How often on average have you eaten 1 fresh banana during the past month?	.62	
						low GHG foods	How often on average have you eaten 1 fresh tomato during the past month?	.89	
						food purchasing	When you buy fresh fruits and vegetables, how often do you make it a point to buy fresh fruits and vegetables that are locally grown?	.71	
						GHG credits	How much of your CO_2 emissions do you offset by buying credits? (single item)	-	
domain-specific measures									
Personal Pro-Environmental Behavior (Walton & Austin, 2011)[b]	personal behavior	residents, USA	6	.76	1		How often do you make an effort to conserve resources in your home, such as electricity, natural gas, and water for environmental reasons?		environmental concern
Organizational Citizenship Behaviour for the Environment (Boiral & Paillé, 2012)[d]	workplace behavior	students, Canada	10		3				
						eco-initiatives	I voluntarily carry out environmental actions and initiatives in my daily work activities.	.92	
						eco-civic engagement	I undertake environmental actions that contribute positively to the image of my organization.	.90	
						eco-helping	I encourage my colleagues to adopt more environmentally conscious behavior.	.81	
Organizational Citizenship Behavior for the Environment (Robertson & Barling, 2017)	workplace behavior	Employees, USA	10	n/p	3				self-reported PEB measure, beliefs that PEBs are inconvenient.
						self-enacted OCBE	At work, I recycle whenever possible.	.80	

						co-worker OCBE	I help my co-workers be environmentally friendly at work.	.88	
						organizational OCBE	I persuade my organization to purchase environmentally friendly products.	.89	
Activism Scale (Séguin, Pelletier, & Hunsley, 1998)[b]	activism	residents, Canada	6	.80	1		How often do you participate in event organized by ecological groups?		perceived importance of environmental problems
Environmental Action Scale (Alisat & Riemer, 2015)	civic actions	students, Canada; MTurkers, USA	18	.92	2				environmental identity, environmental interest, involvement in environmental organizations
						participatory action	I became involved with an environmental group or political party (e.g., volunteer, summer job, etc.).	n/p	
						leadership actions	I organized a community event which focused on environmental awareness.	n/p	
Ecologically Conscious Consumer Behavior (Roberts, 1996)[e]	consumption	adult consumers, USA	22	.96	1		1 have purchased products because they cause less pollution.		perceived consumer effectiveness, environmental concern
Green Consumption (Kim et al., 2012)	consumption	residents, Korea	10	n/p	3				perceived consumer effectiveness, credibility of green products
						health-conscious green consumption behavior	How likely are you to purchase organic foods?	.77	
						resource-conscious green consumption behavior	How likely are you to purchase energy-efficient products?	.74	
						socially conscious green consumption behavior	How likely are you not to purchase products from companies involved with environmental problems?	.92	

Ethically Minded Consumer Behavior (Sudbury-Riley & Kohlbacher, 2016)	consumption	residents > 50 yrs, UK, Germany, Hungary, Japan	10	.86-.93	5				membership in environmental organization, self-reported activism
						ecobuy	I have switched products for environmental reasons.	.78-.90	
						ecoboycott	I do not buy household products that harm the environment.	.85-.90	
						recycle	Whenever possible, I buy products packaged in reusable or recyclable containers.	.80-.91	
						CSRboycott	I will not buy a product if I know that the company that sells it is socially irresponsible.	.83-.90	
						paymore	I have paid more for environmentally friendly products when there is a cheaper alternative.	.91-.96	
Environmentally Responsible Consumption (Gupta & Agrawal, 2018)	consumption	passersby, India	38	n/p	10				environmental group membership, frugality, green consumption value
						purchasing environment-friendly products	I buy products that are environment friendly.	.82	
						need-based purchases	I avoid purchasing things that I do not need.	.86	
						purchasing products in environmentally friendly packaging	I buy products packaged in recyclable material.	.84	
						collaborative consumption	Whenever possible, I borrow things from others.	.82	
						conscious consumption	I avoid wasteful consumption.	.86	
						handling and care	I handle all things with care.	.90	
						repair and reuse	I avoid discarding things that can be repaired.	.89	
						give/donate/offer	I give things that I do not need or use to others.	.90	
						sell/exchange/trade	I sell off things that I do not need or use.	.89	

						disposing of waste	I segregate my household waste before disposing it.	.86	
Pro-Environmental Purchase Behaviour (Tilikidou, Adamson, & Sarmaniotis, 2002)	consumption	residents, Greece	11	.92-.93	1		I try to avoid environmentally harmful products.		other self-reported PEB measure, environmental attitudes
Eco-Socially Conscious Consumer Behavior (Saleem, Eagle, & Low, 2018)	car purchase and use	car dealership customers, Pakistan	9	.81	3				biospheric values
						eco-social conservation	I avoid using wide thread tires for that cause road friction and consume more fuel.	.78	
						eco-social use	Knowing that excessive speed is inefficient and requires more energy to stop the car, I consider observing speed limits.	.74	
						eco-social purchase	I would buy an electric vehicle even if its performance is lower than a conventional car.	.78	
Recycling and Reusing Scales (De Young, 1985-1986)[a,b]	recycling and reusing	residents, USA	11	n/p	2				frugality
						recycling	recycle non-deposit glass jars and bottles	.80	
						reusing	save gift wrapping paper	.84	
Littering Prevention Behavior Scale (Ojedokun, 2016)	littering	residents, Nigeria	41	.81	1		When I see someone littering, I direct him/her politely to use the litter bin.		

Note. [a]No psychometric analysis of the factor structure of the scale is reported.
[b]No psychometric analysis that supported the selection of items is reported.
[c]Additional psychometric evaluation in Arnold et al., 2018; Kaiser & Biel, 2000; Kaiser & Wilson, 2000; Kaiser et al., 2001, 2003, 2018.
[d]Additional psychometric evaluation in Paillé & Boiral, 2013.
[e]Additional psychometric evaluation in Roberts & Bacon, 1997.
*Rasch separation reliability. n/p = not provided.

References (appendix)

Alisat, S., & Riemer, M. (2015). The environmental action scale: Development and psychometric evaluation. *Journal of Environmental Psychology, 43*, 13-23.

Armel, K. C., Yan, K., Todd, A., & Robinson, T. N. (2011). The Stanford Climate Change Behavior Survey (SCCBS): assessing greenhouse gas emissions-related behaviors in individuals and populations. *Climatic Change, 109*, 671-694.

Arnold, O., Kibbe, A., Hartig, T., & Kaiser, F. G. (2018). Capturing the environmental impact of individual lifestyles: evidence of the criterion validity of the general ecological behavior scale. *Environment and Behavior, 50*, 350-372.

Boiral, O., & Paillé, P. (2012). Organizational citizenship behaviour for the environment: Measurement and validation. *Journal of Business Ethics*, *109*, 431-445.

Brick, C., Sherman, D. K., & Kim, H. S. (2017). "Green to be seen" and "brown to keep down": Visibility moderates the effect of identity on pro-environmental behavior. *Journal of Environmental Psychology, 51*, 226-238.

Casey, P. J., & Scott, K. (2006). Environmental concern and behaviour in an Australian sample within an ecocentric–anthropocentric framework. *Australian Journal of Psychology*, *58*, 57-67.

De Young, R. (1985-1986). Encouraging environmentally appropriate behavior: The role of intrinsic motivation. *Journal of Environmental Systems, 15*, 281-292

Erdogan, M., Ok, A., & Marcinkowski, T. J. (2012). Development and validation of children's responsible environmental behavior scale. *Environmental Education Research, 18*, 507-540.

Evans, G. W., Brauchle, G., Haq, A., Stecker, R., Wong, K., & Shapiro, E. (2007). Young children's environmental attitudes and behaviors. *Environment and behavior*, *39*, 635-658.

Gupta, S., & Agrawal, R. (2017). Environmentally Responsible Consumption: Construct Definition, Scale Development, and Validation. *Corporate Social Responsibility and Environmental Management*, *25*, 523–536.

Iwata, O. (2001). Attitudinal determinants of environmentally responsible behavior. *Social Behavior and Personality: an International Journal*, *29*, 183-190.

Kaiser, F. G. (1998). A general measure of ecological behavior. *Journal of Applied Social Psychology*, *28*, 395-422.

Kaiser, F. G., & Biel, A. (2000). Assessing general ecological behavior: A cross-cultural comparison between Switzerland and Sweden. *European Journal of Psychological Assessment*, *16*, 44-52.

Kaiser, F. G., Doka, G., Hofstetter, P., & Ranney, M. A. (2003). Ecological behavior and its environmental consequences: A life cycle assessment of a self-report measure. *Journal of Environmental Psychology*, *23*, 11-20.

Kaiser, F. G., Frick, J., & Stoll-Kleemann, S. (2001). Zur Angemessenheit selbstberichteten Verhaltens: Eine Validitätsuntersuchung der Skala Allgemeinen Ökologischen Verhaltens [Accuracy of self-reports: Validating the general ecological behavior scale]. *Diagnostica*, *47*, 88–95.

Kaiser, F. G., Merten, M., & Wetzel, E. (2018). How do we know we are measuring environmental attitude? Specific objectivity as the formal validation criterion for measures of latent attributes. *Journal of Environmental Psychology*, *55*, 139-146.

Kaiser, F. G., Oerke, B., & Bogner, F. X. (2007). Behavior-based environmental attitude: Development of an instrument for adolescents. *Journal of Environmental Psychology*, *27*, 242-251.

Kaiser, F. G., & Wilson, M. (2000). Assessing People's General Ecological Behavior: A Cross-Cultural Measure. *Journal of Applied Social Psychology*, *30*, 952-978.

Kaiser, F. G., & Wilson, M. (2004). Goal-directed conservation behavior: The specific composition of a general performance. *Personality and Individual Differences, 36,* 1531-1544.

Karp, D. G. (1996). Values and their effect on pro-environmental behavior. *Environment and Behavior, 28,* 111-133.

Kim, S. Y., Yeo, J., Sohn, S. H., Rha, J. Y., Choi, S., Choi, A. Y., & Shin, S. (2012). Toward a composite measure of green consumption: an exploratory study using a Korean sample. *Journal of Family and Economic Issues, 33,* 199-214.

Larson, L. R., Stedman, R. C., Cooper, C. B., & Decker, D. J. (2015). Understanding the multi-dimensional structure of pro-environmental behavior. *Journal of Environmental Psychology, 43,* 112-124.

Lee, T. H., Jan, F. H., & Yang, C. C. (2013). Conceptualizing and measuring environmentally responsible behaviors from the perspective of community-based tourists. *Tourism Management, 36,* 454-468.

Lee, Y. K., Kim, S., Kim, M. S., & Choi, J. G. (2014). Antecedents and interrelationships of three types of pro-environmental behavior. *Journal of Business Research, 67,* 2097-2105.

Maloney, M. P., & Ward, M. P. (1973). Ecology: Let's hear from the people: An objective scale for the measurement of ecological attitudes and knowledge. *American Psychologist, 28,* 583-586.

Maloney, M. P., Ward, M. P., & Braucht, G. N. (1975). A revised scale for the measurement of ecological attitudes and knowledge. *American Psychologist, 30,* 787-790.

Markle, G. L. (2013). Pro-environmental behavior: Does it matter how it's measured? Development and validation of the Pro-Environmental Behavior Scale (PEBS). *Human Ecology, 41,* 905-914.

Ojedokun, O. (2016). Development and Psychometric Evaluation of the Littering Prevention Behavior Scale. *Ecopsychology, 8*, 138-152.

Roberts, J. A. (1996). Green consumers in the 1990s: profile and implications for advertising. *Journal of Business Research, 36*, 217-231.

Robertson, J. L., & Barling, J. (2017). Toward a new measure of organizational environmental citizenship behavior. *Journal of Business Research, 75*, 57-66.

Saleem, M. A., Eagle, L., & Low, D. (2018). Climate change behaviors related to purchase and use of personal cars: Development and validation of eco-socially conscious consumer behavior scale. *Transportation Research Part D: Transport and Environment, 59*, 68-85.

Schultz, P. W., Gouveia, V. V., Cameron, L. D., Tankha, G., Schmuck, P., & Franěk, M. (2005). Values and their relationship to environmental concern and conservation behavior. *Journal of Cross-Cultural Psychology, 36*, 457-475.

Séguin, C., Pelletier, L. G., & Hunsley, J. (1998). Toward a model of environmental activism. *Environment and Behavior, 30*, 628-652.

Sia, A. P., Hungerford, H. R., & Tomera, A. N. (1986). Selected predictors of responsible environmental behavior: An analysis. *The Journal of Environmental Education, 17*, 31-40.

Stanley, L. R., Lasonde, K. M., & Weiss, J. (1996). The relationship between environmental issue involvement and environmentally-conscious behavior: An exploratory study. *Advances in Consumer Research, 23*, 183-188.

Stern, P. C., Dietz, T., Abel, T., Guagnano, G. A., & Kalof, L. (1999). A value-belief-norm theory of support for social movements: The case of environmentalism. *Human Ecology Review*, 81-97.

Stern, P. C. (2000). New environmental theories: toward a coherent theory of environmentally significant behavior. *Journal of Social Issues, 56*, 407-424.

Stone, G., Barnes, J. H., & Montgomery, C. (1995). Ecoscale: a scale for the measurement of environmentally responsible consumers. *Psychology & Marketing, 12*, 595-612.

Sudbury-Riley, L., & Kohlbacher, F. (2016). Ethically minded consumer behavior: Scale review, development, and validation. *Journal of Business Research, 69*, 2697-2710.

Tilikidou, I., Adamson, I., & Sarmaniotis, C. (2002). The measurement instrument of ecologically conscious consumer behaviour. *MEDIT, 1*, 46-53.

Walton, T., & Austin, D. M. (2011). Pro-environmental behavior in an urban social structural context. *Sociological Spectrum, 31*, 260-287.