Accepted for Publication in the Journal of Environmental Psychology

Word count: 8,617

Offsetting behavioral costs with personal attitude:

Identifying the psychological essence of an environmental attitude measure

Florian G. Kaiser and Florian Lange

Otto-von-Guericke University Magdeburg, Germany

Katholieke Universiteit Leuven, Belgium

Author Note

This research was financially supported by the Fonds Wetenschappelijk Onderzoek – Vlaanderen (FWO) and the European Union Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie Grant agreement No 665501. We thank Juliane Bücken for helping with our statistical analyses, Ronja Gerdes for double-checking the numerical information, Laura Henn, Anne Overbeck, Mark Wilson, as well as the editor, Mark Ferguson, and three anonymous reviewers for their helpful comments on earlier versions of this article, and Jane Zagorski for her language support.

Correspondence concerning this article should be addressed to Florian Kaiser, Otto-von-Guericke University, Institute of Psychology, P.O. Box 4120, D-39016 Magdeburg, Germany; e-mail: florian.kaiser@ovgu.de

Abstract

Identifying the essence of a latent psychological attribute captured by a measurement instrument requires more than a glance at the measure's indicators or confidence in an established measurement practice. It demands an exploration of the nomological network in which an attribute (e.g., environmental attitude) is theoretically anticipated to operate. With our study, we aimed to identify the General Ecological Behavior (GEB) scale as a Campbell-paradigm-based measure of environmental attitude. In our empirical test ($N = 183$), we juxtaposed the two estimated parameters—behavioral costs and what is believed to represent environmental attitude—with *real* costs. To do so, costs were manipulated as objective waiting times that differentiated between otherwise identical pro-environmental behaviors. Our findings corroborate the idea that the GEB's cost estimates reflect actual behavioral costs. Furthermore, progressively increasing costs were the expected progressive impediment of behavior that, according to Campbell's paradigm, environmental attitude is presumed to offset.

Keywords: attitude measurement; environmental attitudes; attitude-behavior consistency; conservation (ecological behavior); Campbell paradigm

1. Introduction

Sustainably changing people's consumptive ways of life is possible when we, in environmental protection research, can identify psychological attributes (e.g., environmental attitude) that truly help to promote environmentally protective engagement. Unfortunately, psychological attributes are typically not directly observable. Identifying individual differences in such latent attributes faces an intractable problem. It requires knowledge about an attribute's indicators, and knowledge about these indicators for its part requires knowledge about the attribute (De Houwer et al., 2013). Therefore, the epistemological question is: How can we know we are measuring environmental attitude and not some other latent attribute (e.g., environmental identity or conformity)?

To gain knowledge about the psychological essence of an attribute, which cannot be observed directly, empirical scientists need to explore the nomological network (i.e., the theoretically expected connections) in which an attribute is anticipated to operate. If some of the network connections involve observables, empirically identifying the psychological essence becomes possible (see Cronbach & Meehl, 1955).¹ Examples of such observables are clinical symptoms, corporal reactions of sorts, tangible consequences of behavior (e.g., ecological footprints or amounts of energy consumed), and, of course, manifest behavior and manifest costs of behavior. Ideally, identifying a measurement instrument as a measure of a specific latent attribute will be done experimentally, using observables from outside the measurement process itself (Whitely, 1977).

In our research, we derived an experimental test from the nomological network (the Campbell paradigm) in which the latent psychological attribute of interest (environmental attitude) was expected to operate. We employed an observable that is distinct from the actual measurement instrument: the waiting time people incur by choosing to ride a bicycle (the behavioral alternative with the smaller energy expenditure). Waiting time is meant to reflect

an impediment to behavior that environmental attitude is expected to offset before the behavior has a reasonable chance of manifesting itself.

We begin by illustrating the problem that plagues the measurement of unobservable attributes. In the subsequent section, we describe the specific nomological network that illustrates *how* environmental attitude (the latent psychological attribute) and manifest costs of pro-environmental behavior are interrelated: the Campbell paradigm. Finally, we detail the experimental test with which we aimed to identify the General Ecological Behavior (GEB) scale—a well-established instrument that is based on self-reports of pro-environmental behaviors (see Lange & Dewitte, 2019b)—as a Campbell-paradigm-based measure of environmental attitude.

1.1. Self-reports of pro-environmental behaviors

Because latent attributes can become objectively observable only indirectly in people's corporal movements and reactions (De Houwer et al., 2013), latent attributes are typically inferred from behavior: either from verbal behavior (i.e., self-reports of opinions, judgments, appraisals, affect, or past behavior) or from nonverbal behavior observed in labs or in real-life situations (Kaiser et al., 2018; Kaiser & Wilson, 2019). Such behavioral indicators that are used to measure an attribute are, in return, also used to *operationally* define the essence of a measurement instrument. For instance, self-reports of past environmentally protective behaviors are often regarded as measures of pro-environmental behaviors (see, e.g., Kormos & Gifford, 2014; Lange & Dewitte, 2019b), and evaluative statements and personal appraisals specify environmental attitude measures (see, e.g., Milfont & Duckitt, 2004; Schultz, 2001).

Operationally defined measurement instruments—although widely common in psychology—are insufficient because they are grounded in a circular logic (i.e., they are tautological): they already require knowledge about the measurement of an attribute that such a measure is created to accrue (De Houwer et al., 2013). Operationally defined measures become even more questionable when they are later found not to operate with observables in

the ways they were theoretically anticipated to operate (see Cronbach & Meehl, 1955). We believe this is the case when environmental attitude—grounded in traditional measures—seems unable to explain manifest pro-environmental behaviors (see, e.g., Gifford, 2014; Kollmuss & Agyeman, 2002; Stern, 2000a).

Fairly recently and in contrast to common practice, Kaiser and colleagues have come to reinterpret their measure of self-reports of past environmentally protective behavior (the General Ecological Behavior [GEB] scale; e.g., Kaiser, 1998; Kaiser & Wilson, 2000, 2004) as a measure of environmental attitude (see, e.g., Kaiser et al., 2007, 2010, 2013). In several experiments and in contrast to studies employing more traditional measures (see, e.g., Bickman, 1972), their compound measure of “environmental attitude,” consisting of behavioral self-reports, was found to reliably account for manifest pro-environmental behavior: the demand for resources in a commons dilemma game (Kaiser & Byrka, 2015), the proportion of vegetarian lunches (Kaiser et al., 2020), the choice of CO₂-optimal routes (Taube et al., 2018), and the proportion of green products (Taube & Vetter, 2019).

Before explaining how we devised a rigorous experimental trial to test the GEB as a measure of environmental attitude, we will first provide a brief account of the nomological network (i.e., the Campbell paradigm) that was proposed by Kaiser et al. (2010) as their measurement theory. The paradigm is named after Donald T. Campbell (1963), who had originally suggested that the costs of behavior could be used to measure attitudes.

1.2. The Campbell paradigm

The nomological network—the Campbell paradigm—is a variant of Rosenberg and Hovland’s (1960) generic measurement theory of attitudes (see Kaiser & Wilson, 2019), which also inspired what was probably the first environmental attitude measure ever (see Maloney & Ward, 1973; Maloney et al., 1975). According to the Campbell paradigm, people disclose their attitude levels—their valuations of the attitude object (e.g., environmental protection) or their commitment to the attitude-implied behavioral goal (e.g., protecting the

environment)—in the extent to which they engage in verbal and nonverbal behaviors that involve increasing levels of behavioral costs (see Kaiser, in press; Kaiser et al., 2010; Kaiser & Wilson, 2019). In other words, people's levels of environmental attitude can be derived from the environmentally protective activities they engage in (see, e.g., Kaiser et al., 2007, 2013). For instance, when people publicly express approval for environmental protection, recycle paper and cardboard, refrain from commuting by car, and eliminate meat from their diets, they reveal their personal commitment to protecting the environment.²

Any behavior aimed at environmental protection involves costs. Behavioral costs can be small (e.g., when a person verbally expresses approval for environmental protection), or they can be substantial (e.g., when a person commutes by bicycle in inclement weather). Irrespective of the specific attitude level, people disclose their environmental attitude in a cost-effective, rational manner (e.g., Kaiser et al., 2010, 2013). In other words, people generally favor more convenient and socially accepted over more demanding, socially prohibited, or otherwise costly environmentally protective behaviors. For example, people are more likely to express environmental concern in surveys (something that is easy to do) than to actively engage in exhausting practices (e.g., commuting by bicycle).

Engagement in a specific behavior involves costs in terms of time, money, effort, courage, inconvenience, etcetera. Such costs have a chance of being endured by people only when these people's levels of environmental attitude at least match the costs (see Campbell, 1963; Kaiser, in press; Kaiser et al., 2010, and Figure 1). Consequently, people who engage in a comparatively demanding environmentally protective behavior (e.g., became members in environmental organizations) reveal higher levels of environmental attitude than people who fail to engage in that behavior (e.g., nonmembers; see Arnold & Kaiser, 2018).

Note that the highly parsimonious understanding of manifest engagement in a behavior depicted in Figure 1 with only two *unconditionally* effective influences differs markedly from other more complex behavior models in environmental protection research

(see, e.g., Bamberg & Möser, 2007; Klöckner & Blöbaum, 2010; Stern, 2000b) as they typically involve more than two foundations of behavior. It also differs from models that condition the behavioral efficacy of environmental attitude on behavioral costs (see, e.g., Diekmann & Preisendörfer, 1998; Schultz & Oskamp, 1996). The model in Figure 1 specifically implies that engagement in a behavior occurs if people's attitude levels are powerful enough to offset the specific costs of the particular behavior (Kaiser, in press).³

Figure 1. The two critical foundations of pro-environmental behavior (i.e., behavior aimed at environmental protection) according to the Campbell paradigm.

Mathematically, the Rasch model describes the theoretically anticipated connection depicted in Figure 1 between a person's environmental attitude, the person's likelihood of engaging in a particular environmentally protective activity, and the cost of engaging in that behavior (for more details about the Rasch model, see Rasch, 1960/1980; for a more recent account, see, e.g., Wilson, 2005):

$$\ln\left(\frac{p_{ki}}{1-p_{ki}}\right) = \theta_k - \delta_i \quad (1)$$

In this equation, the natural logarithm of the ratio of the probability (p_{ki}) of person k 's engagement relative to the probability of non-engagement ($1-p_{ki}$) in a specific environmentally protective activity i (i.e., its odds) is given by the difference between the estimate of k 's environmental attitude (θ_k) and the estimate of the costs of behavior i (δ_i). This compensatory relation between the attitude and cost estimates basically implies that an individual's environmental attitude needs to clear the costs of a behavior before the behavior has a reasonable chance ($p_{ki} \geq .50$) of being implemented by the person. In other words, an

individual's attitude needs to offset the costs of a behavior before a behavior can become manifest.

1.3. Research goals

Our goal was to identify the General Ecological Behavior (GEB) scale as a Campbellian environmental attitude measure. To do so, we employed the Pro-Environmental Behavior Task (PEBT) by Lange et al. (2018). The PEBT involves manifest behavior with real pro-environmental consequences (albeit small ones in absolute terms). More importantly though, it employs objective impediments (i.e., behavioral costs) that come with acting pro-environmentally: waiting times. Because people's protecting the environment in the PEBT comes with real costs, the PEBT allowed us to experimentally manipulate costs outside the measurement process and to link them, for the first time, with the costs of behavior and with people's attitude levels that were both estimated within the measurement process.

In the nomological network of our measurement theory, the Campbell paradigm, cost estimates represent behavioral costs. Thus, we expected that the objective costs of a behavior, manipulated as different waiting times, would correspond with the behavioral cost estimates that derived from the Rasch scale calibration (*Hypothesis 1*). In addition, in this paradigm, in order for behavior to have a reasonable chance of becoming manifest, an environmental attitude has to offset the objective behavioral costs. Hence, we expected that enduring increasing waiting times in the PEBT would correspond with increasing environmental attitude estimates as well (*Hypothesis 2*). In this manner, by using observables (i.e., real waiting times) defined and manipulated outside the measurement process, we directly tested what Wilson (2005) called the internal structure of the measurement theory of environmental attitude, the Campbell paradigm (see, e.g., Kaiser et al., 2010).

2. Method

Our research adheres to the American Psychological Association's (APA, 2017) ethical principles and ethical code of conduct for research with human participants. The

current data were collected as part of a larger study (see Lange & Dewitte, 2020) that was approved by the Social and Societal Ethics Committee at Katholieke Universiteit Leuven (G-2016 12 694). Before we implemented the procedures outlined below, we asked participants to view and rate a number of videos that differed in affective valence (neutral vs. positive) between experimental groups. The experimental groups were not significantly different in their estimated average environmental attitude levels. Thus, a confounding influence of video valence seemed unlikely. The complete study protocol and all data can be found at osf.io/cwyze/.

2.1. Participants

Participants were recruited from the international subject pool of the Faculty of Economics and Business at Katholieke Universiteit Leuven, Belgium. In the invitation email, potential participants were informed that they needed to have a proficient grasp of English. Data were collected from $N = 183$ volunteers. The sample size was based on power considerations related to the anticipated size of the effect from the affect videos in the larger study (i.e., $d = 0.50$, to be detected with 95% power using a one-sided t test at $\alpha = .05$). In general, sample sizes larger than $N = 100$ can be expected to yield reasonably precise cost estimates within a Rasch scale calibration (see Wright, 1977).

The ages of the final sample of participants (60 men, 123 women) ranged from 18 to 64 ($M = 24.15$, $SD = 6.26$). Most participants (86%) indicated that they were students. The remaining participants were employed full-time (10%), were employed part-time (2%), were unemployed (1%), or preferred not to reveal their occupational status (1%). Participants were tested in groups of up to six and received € 10 (approximately \$12 US) for their participation.

2.2. Procedure and measures

Upon arriving at the laboratory, participants were asked to provide informed consent before they were shown to one of six partially enclosed cubicles where they completed the study in private. After the video-rating task (unrelated to the present research question), they

completed a 24-trial version of the Pro-Environmental Behavior Task (PEBT) before being directed to a Qualtrics questionnaire. Among other scales, the questionnaire included the proposed measure of environmental attitude.

Environmental attitude. As a measure of environmental attitude, the General Ecological Behavior (GEB) scale was calibrated in its typical way (see, e.g., Kaiser & Byrka, 2015; Taube & Vetter, 2019) as a Rasch scale using the dichotomous Rasch model (see Rasch, 1960/1980). The instrument consisted of 50 self-reports, such as "I am a member of an environmental organization" and "I collect and recycle used paper" (for a complete list of the items, their cost estimates, and their fit statistics, see Table 1A in the Appendix). The self-reports used in the present research were slightly amended from the instrument proposed by Kaiser and Wilson (2004). In numerous previous studies, the instrument has been shown to be reliable (for an overview of its psychometric features, see Kaiser et al., 2010; Lange & Dewitte, 2019b).

Participants responded to each specific behavior using either a 5-point frequency format (*never* to *very often*) or a *yes-no* format. *Not applicable* was a response alternative that was available when participants were unable to provide an answer. The *not applicable* responses were treated as missing values. Note that Rasch scale calibrations and, thus, parameter estimations can be achieved even with incomplete data records that contain missing values as the estimation procedure is based on a maximum likelihood approach (for more details, see, e.g., Embretson & Reise, 2000).⁴ Negatively formulated self-reports ($i = 19$), which represented environmentally harmful behavior, were inversely coded.⁵

In their research with the GEB, and contrary to common expectations, Kaiser and Wilson (2000) found that binary responses were superior to graded responses (which were thought to capture increasing frequencies of engagement). As participants used these graded responses inconsistently, the frequency order seemed to reflect personal idiosyncrasies (i.e., measurement error) rather than valid differences in people's levels of environmental attitude.

The same was the case for the present study.⁶ In view of these findings and in line with the established practice when using the GEB, we decided to once again recode the 32 items that employed the frequency response format into a binary format, with the responses *never*, *seldom*, and *occasionally* recoded as environmentally harmful behavior. The response options *often* and *very often* were recoded as pro-environmental behavior.

Note that this practice of dichotomization is well-established and well-justified, particularly in attitude research, for the very reason we use it as well: to prevent measurement error (see DeCoster et al., 2009; for supporting evidence, see, e.g., Matell & Jacoby 1971; Peabody, 1962; for a recent example with another environmental attitude measure, see Zhu & Lu, 2017, p. 88). Note also that the proposed dichotomization practice in which we limit the number of response options in items as they are administered in the survey (to produce more valid and reliable answers by eliminating error variance) *before* calibrating a measure must be distinguished from the indefensible practice of obliterating systematic variance in an established measure *after* calibration (see, e.g., DeCoster et al., 2009; MacCallum et al., 2002). Accordingly, we argue that dichotomizing items before calibrating a measurement instrument is a sensible way to deal with item-level measurement error.

Environmental attitude was assessed in logits. Logits stand for the natural logarithm of the ratio of the probability of engagement to nonengagement in the pro-environmental self-reports. Larger positive logits reflect a more pronounced environmental attitude. The GEB's ability to distinguish individuals according to their levels of environmental attitude was reasonable (person separation reliability: $rel = .76$). Fit values for the self-reports, which reflect the relative discrepancy between the model predictions and the observed data, showed rather acceptable model fit with mean square (MS) values weighted by the item variance that fell within $MS \geq 0.80$ and $MS \leq 1.20$ (see Wright et al., 1994; for more details regarding global and item-specific fit statistics, see the *basic* model in Table 1 and Table 1A in the Appendix).

The Pro-Environmental Behavior Task (PEBT). With this task, people's environmental protection engagement (i.e., their *manifest* pro-environmental behavior) was tested (see Lange et al., 2018; Lange & Dewitte, 2019a). Every PEBT trial consisted of a choice between an environmentally harmful option (labeled *car*) and a pro-environmental option (labeled *bicycle*).

Choosing the bicycle option was related to actual behavioral costs because participants had to wait longer (compared with the car option) before they could move on to the following trial. The extra waiting times incurred from choosing the bicycle option varied as a within-subject factor at levels of 5, 10, 15, 20, 30, or 60 s. As such, these costs (i.e., the waiting times that distinguished six otherwise identical pro-environmental behaviors) were manipulated outside the measurement process (i.e., the parameter estimation procedure).

Each of the six waiting times was presented four times to each participant, thus yielding a total of 24 trials. The trial order was randomized with a constraint (each waiting time difference could only be presented for the $n+1^{\text{st}}$ time when all waiting time differences had been presented n times). Participants who consistently chose the bicycle option prolonged their experimental session by more than 9 min (with no additional financial compensation) compared with participants who consistently chose the car option.

Per waiting time, participants thereby had four opportunities to act pro-environmentally. Accordingly, people's manifest pro-environmental engagement could have a score between 0 (exclusively environmentally harmful car choices, zero pro-environmental bicycle choices) and 4 (exclusively pro-environmental bicycle choices). As with the GEB items, we recoded the PEBT items to prevent erratic response patterns.⁷ Consequently, we had to be specific about what we considered to be pro-environmental behavior on the PEBT. As we did not wish to distort our findings, we applied two interpretations of pro-environmental engagement.

To capture a lenient view, participants behaved pro-environmentally when they chose the pro-environmental bicycle option on at least one of the four trials with the same waiting time (e.g., 30 s). Exclusively choosing the car was coded 0 (environmentally harmful behavior), choosing the bicycle option on at least one trial was coded 1 (pro-environmental behavior). By contrast, the strict view defined pro-environmental behavior as exclusively choosing the pro-environmental bicycle option. Scores of 0 to 3 were all coded 0 (environmentally harmful behavior), whereas only a score of 4 was coded 1 (pro-environmental behavior).

In addition to the behavioral costs, participants' choices on the PEBT had actual consequences for the environment in terms of energy consumption and associated CO₂ emissions. When participants chose the car option, 12 USB-powered lights, each containing eight light-emitting diodes (LEDs), were turned on. Lights were placed under participants' desks to limit distraction during the study. Each light measured approximately 3 cm x 3 cm x 3 cm and emitted white light (more details about the setup can be found in Lange et al., 2018). Note that these environmental consequences made choosing the bicycle option objectively pro-environmental. Lights were illuminated during the waiting period following the car choice. During the waiting period, a display was presented to inform participants of the status of the lights (i.e., on or off).

On each trial, participants received prospective information about (a) the travel times associated with the two options, (b) the extension of the travel time when they chose the bicycle option, (c) the number of lights illuminated when they chose the car option, and (d) the associated amount of CO₂ emissions. Before the beginning of the task, they received general instructions about the task contingencies (see Lange et al., 2018) and completed two practice trials to familiarize themselves with these contingencies.

3. Results

Results are reported in two sections. In the first section, we tested the measurement theory as represented in the Campbell paradigm, which we employed to measure environmental attitude. To do so, we jointly calibrated the 50 behavioral self-reports from the General Ecological Behavior (GEB) scale—the presumed Campbellian environmental attitude measure—with the six manifest behaviors from the Pro-Environmental Behavior Task (PEBT). Successful scale calibration represents a rigorous empirical test of the formal features of the Campbell paradigm, and, as such, it is the precondition for testing the two Campbell-paradigm-related hypotheses in the second section. There, we detail what we found when we connected the cost and attitude estimates derived from the measurement process with the objective costs (i.e., waiting times) of pro-environmental behavior on the PEBT.

3.1. Instrument calibration with two views of pro-environmental engagement

Next to the GEB alone, we jointly calibrated the 50 GEB items and the six manifest pro-environmental behaviors (with varying waiting times) from the PEBT using the dichotomous Rasch model. Separate calibrations were performed for the two views (i.e., lenient or strict) of manifest pro-environmental behavior in the PEBT. As can be derived from Table 1, the psychometric features of all three instruments—the basic environmental attitude scale consisting of the 50 behavioral self-reports of the GEB and the lenient and strict variants of the same scale consisting of the 50 self-reports plus the six PEBT behaviors—were rather comparable in their overall fit statistics (for item-specific statistics, see Table 1A in the Appendix). Moreover, the three estimates of a person's environmental attitude were comparable too. Irrespective of what instrument was employed, we ended up with three virtually indistinguishable sequences of persons as can be seen in the Pearson correlations between the attitude estimates of the three instruments: $r_{\text{basic-lenient}} = .97$, $r_{\text{basic-strict}} = .94$, $r_{\text{lenient-strict}} = .97$ (all three coefficients presented *without* correcting for measurement error attenuation).

Table 1

Psychometric Features of Three Environmental Attitude Measures

	Basic model	Lenient model	Strict model
ITEM FIT STATISTICS:			
<i>M(MS)</i>	1.00	1.00	1.00
<i>SD(MS)</i>	0.08	0.08	0.08
<i>Minimum (MS)</i>	0.86	0.86	0.88
<i>Maximum (MS)</i>	1.25	1.25	1.26
PERSON FIT STATISTICS:			
<i>M(MS)</i>	1.00	1.00	1.00
<i>SD(MS)</i>	0.24	0.24	0.23
<i>% people with poor fit (t ≥ 1.96)</i>	6.01%	5.46%	7.10%
<i>Separation reliability</i>	.76	.77	.77

Note. The *basic* model was based on 50 binary-coded behavioral self-reports. The other two models were based on the same 50 binary-coded behavioral self-reports plus the six binary-coded manifest pro-environmental behaviors from the PEBT. Whereas the *lenient* model describes pro-environmental engagement as a minimum of one bicycle choice out of four such opportunities with a specific waiting time, the *strict* model describes pro-environmental engagement as exclusively choosing the bicycle (i.e., four times) with a specific waiting time. Mean square (*MS*) and *t* values were weighted by the item variance. They reflect the relative discrepancy between the model predictions and the observed data and were used to assess model-data fit (see, e.g., Wright & Masters, 1982).

In summary, including the manifest pro-environmental behaviors from the PEBT had only a marginal effect on the calibration of the scale. The person fit statistics improved slightly for the lenient view and declined slightly for the strict view of pro-environmental behavior. The reliability improved minimally in both instances. From these data, we concluded that including the manifest pro-environmental behaviors from the PEBT—irrespective of which view of pro-environmental behavior was chosen (i.e., lenient or strict)—

in the Campbell-paradigm-based instrument was possible without any substantial changes in the psychometric features of the presumed environmental attitude measure.

Integrating the six manifest pro-environmental behaviors from the PEBT into the original Campbell-paradigm-based measure (i.e., the GEB) was critical for achieving our ultimate objectives. In the next section, we describe how we tested our two hypotheses in which we linked the objective behavioral costs (defined outside the measurement process as different waiting times in the PEBT) with the internal structure of the Campbell paradigm.

3.2. Hypothesis tests

The Campbell paradigm demands that the costs of behavior, the *increasing waiting times* that differentiate otherwise identical pro-environmental behaviors in the PEBT, correspond with *the behavioral costs* that were estimated in the measurement process (*Hypothesis 1*). Moreover, progressive waiting times (i.e., progressive *costs of behavior*, defined outside the measurement process) must be matched with progressive individual *environmental attitude levels* estimated in the measurement process (*Hypothesis 2*). Note that support for either one of the two hypotheses qualifies as an empirical finding and does not represent an artifact of the Rasch model calibration (for supporting evidence, see Table 3S in the Supplementary data).

The first hypothesis concerns the question of whether increasing cost estimates validly reflect increasing actual waiting times associated with the pro-environmental bicycle choice (i.e., 5, 10, 15, 20, 30, or 60 s) as implied by the Campbell paradigm. As can be seen in Figure 2, we found that the cost estimates mimicked increasing waiting times (in s) without exception. In and of itself, this already represents circumstantial evidence in support of the corresponding claim implied by the Campbell paradigm, namely, that different waiting times represent differentially effective behavioral costs.

To conclusively corroborate the correspondence between waiting times and cost estimates, it was also necessary to test whether objectively different waiting times resulted in

significantly different cost estimates. By comparing cost estimates for adjacent waiting times, we could statistically test the sensitivity of cost estimates for differences in actual waiting times between otherwise identical pro-environmental behaviors.

Figure 2. Persons clustered at environmental attitude levels, GEB self-reports clustered at cost estimates, their respective frequencies, and the six PEBT behaviors with different waiting times depicted at the level of their estimated costs (with either a lenient or a strict view of these pro-environmental behaviors). All numerical information has been converted into the metric of the lenient instrument to make the numbers directly comparable (and, accordingly, presentable in a single figure). Numerical information regarding costs for the lenient view of the PEBT items is depicted in Table 2a; for the strict view of the PEBT items, see Table 2b.

Table 2 shows that nine of the 10 comparisons of pairs of behavioral cost estimates were significant in the expected direction, reflecting the idea that longer waiting times rendered engagement less likely to occur: seven at $p \leq .01$ —apparent in the nonoverlapping

95% confidence intervals—and two at $p \leq .05$ (the lightly shaded comparisons in Table 2)—apparent in an overlap that was half or less than half of the width of the 95% confidence interval (cf. Cumming & Finch, 2005). Only the 10 s waiting time was not more challenging for people than the 5 s waiting time when environmentally protective engagement was leniently interpreted (see the darkly shaded comparisons in Table 2a). In other words, irrespective of people's attitude levels, it was not more demanding to endure a 10 s waiting time at least once than it was to endure a 5 s waiting time at least once. Note that finding at least nine significant ($p < .05$) differences in 10 statistical tests is, according to a binomial test, statistically unlikely ($p < .001$).

Table 2

Estimated Costs of Environmental Behaviors with Distinct Waiting Times (in s)

(a) a minimum of one pro-environmental (i.e., bicycle) choice

WAITING TIME FOR BICYCLE CHOICES:	δ_{shorter}	Upper bound	Lower bound	δ_{longer}
30 s vs. 60 s	-0.39	-0.08	0.29	0.60
20 s vs. 30 s	-1.12	-0.75	-0.70	-0.39
15 s vs. 20 s	-2.11	-1.62	-1.49	-1.12
10 s vs. 15 s	-3.49	-2.59	-2.60	-2.11
5 s vs. 10 s	-5.10	-3.14	-4.39	-3.49

(b) exclusively pro-environmental (i.e., bicycle) choices

WAITING TIME FOR BICYCLE CHOICES:	δ_{shorter}	Upper bound	Lower bound	δ_{longer}
30 s vs. 60 s	1.26	1.59	1.67	2.11
20 s vs. 30 s	0.73	1.04	0.93	1.26
15 s vs. 20 s	0.09	0.40	0.42	0.73
10 s vs. 15 s	-0.92	-0.57	-0.22	0.09
5 s vs. 10 s	-2.04	-1.55	-1.27	-0.92

Note. All numerical information was converted into the logit metric of the lenient instrument to make the numbers directly comparable. Comparisons of cost estimates (δ) of environmental behaviors with adjacent waiting times are presented. Behavioral cost estimates are expressed in logits for behaviors with shorter (δ_{shorter}) and longer (δ_{longer}) waiting times. The upper bound represents the upper end of the 95% confidence interval of the cost estimate of the shorter

waiting time (e.g., 30 s). The lower bound represents the lower end of the 95% confidence interval of the cost estimate of the longer waiting time (e.g., 60 s). Nonshaded comparisons represent significantly distinct (i.e., nonoverlapping) cost estimates at $p \leq .01$ (i.e., when the upper bound figure is smaller than its counterpart lower bound figure). Lightly shaded comparisons represent cost estimates that were also statistically different at $p \leq .05$. Darkly shaded comparisons represent cost estimates that were not statistically different. Panel a provides cost estimates of the lenient interpretation of pro-environmental behavior and Panel b of the strict interpretation.

In summary, of the 10 tests, nine corroborated the expected cost-estimate increase that followed the objective increase in PEBT waiting times. Only the two shortest waiting times did not differentially constrain the occasional (one in four opportunities) engagement in a pro-environmental behavior.

The different views of pro-environmental behavior (i.e., strict vs. lenient) resulted in the anticipated differences in the cost estimates of the six PEBT behaviors. Apparently, the recoding of the PEBT behaviors did not erratically affect the cost estimates as the lenient cost estimates could be linearly translated into the strict cost estimates with virtually no loss ($R^2 = .999$): $\delta_{\text{strict}} = .72 * \delta_{\text{lenient}} + 1.59$ (see Figure 3). In other words, cost estimates increased significantly when behavior was defined strictly (i.e., exclusive pro-environmental choices) in comparison with when behavior was defined leniently (i.e., minimally one pro-environmental choice; for details, see Table 2A in the Appendix). This systematic difference also became apparent in the respective cost estimates of the six PEBT behaviors in Figure 2. Expectedly, a strict view of pro-environmental behavior rendered engagement generally more demanding in comparison with a more lenient view.

As surmounting behavioral costs was in turn expected to uncover people's levels of environmental attitude, individuals with higher levels of environmental attitude were expected to offset more extensive costs—and engage in objectively more demanding pro-environmental behaviors (*Hypothesis 2*). We tested this hypothesis by connecting the

objective costs of the waiting times that people endured in order to engage in a pro-environmental activity (i.e., accepting the 5, 10, 15, 20, 30, or 60 s waiting times) with estimates of the level of environmental attitude required to do so.

Figure 3. The six PEBT items, representing pro-environmental behaviors with different waiting times (5, 10, 15, 20, 30, 60 s), each represented with two cost estimates, one for the strict view (i.e., exclusively choosing the bicycle) and one for the lenient view (i.e., a minimum of one bicycle choice). All numerical information has been converted into the metric of the lenient instrument.

Irrespective of whether we applied the lenient or strict view of pro-environmental behavior, we found that the presumed measure of environmental attitude was systematically linked with the costs people endured to act pro-environmentally. With increasingly more strenuous environmentally protective activities (i.e., with increasing waiting time costs associated with choosing the bicycle on the PEBT), the environmental attitude that was required for a behavior to have a reasonable chance ($p = .50$) to become manifest had to increase as well (see Figure 2). This result is in line with the idea depicted in Figure 1 that increasing costs have to be offset by increasing levels of attitude.

What becomes apparent from the correspondence between the six PEBT behaviors and people's attitude levels in Figure 2 is that a lenient understanding of pro-environmental engagement renders part of the PEBT behaviors overly undemanding. Even participants with low levels of environmental attitude were quite ready to act pro-environmentally at least once when it extended their waiting time by only 5 or 10 s. This can also be seen in the increased uncertainty of the cost estimates of the two easiest pro-environmental behaviors from the PEBT (evident in their increased confidence intervals when compared with the confidence intervals of the cost estimates of the other pro-environmental behaviors in Table 2).

In summary, environmental attitude reflects the behavioral costs people actually overcome by implementing pro-environmental behaviors. However, this is only possible when the objective costs of pro-environmental behavior are demanding enough to discriminate across people with different environmental attitude levels.

4. Discussion

In our research, we integrated the six manifest pro-environmental behaviors from the Pro-Environmental Behavior Task (PEBT; see Lange et al., 2018) into the Campbell-paradigm-based instrument—the General Ecological Behavior (GEB) scale—meant to measure environmental attitude (see Figure 2). This gives unique credit to the notion that manifest pro-environmental behaviors operate in a manner that is comparable to self-reports of pro-environmental behavior as indicators of environmental attitude without any substantial deterioration in the psychometric features (i.e., reliability and fit statistics) of the measurement instrument (see Table 1). This is the case irrespective of the particular—lenient or strict—definition of pro-environmental behavior. Note that the possibility that both self-reports and manifest environmentally protective behaviors can be used in the measurement of environmental attitude by no means implies that self-reports also work as measures of manifest pro-environmental behavior. (Several colleagues would righteously disagree with the latter conclusion: see, e.g., Kormos & Gifford, 2014; Lange & Dewitte, 2019b.)

Our finding adds to prior research indicating that individual levels of an attitude can be derived from all kinds of verbal acts—specifically, self-reports of past behavior and evaluative statements—addressing one and the same attitudinal object (e.g., environmental protection; see, e.g., Kaiser et al., 2018). Evidentially, environmental attitude can also be derived from manifest behavior and behavioral self-reports as long as their differential behavioral costs are considered in the particular measurement model (by employing a model from the family of Rasch models). As such, our findings can be contrasted against traditional views in environmental psychology that speak of the essence of measures to be connected with the types of indicators that are used. Rather, manifest behaviors and self-reports of behavior, as well as verbal expressions of appreciation and opinions, represent suitable indicators of environmental attitude (see Kaiser & Wilson, 2019; Kaiser et al., 2010, 2018; but also Rosenberg & Hovland, 1960).

Adding to prior experimental research that tested the compensatory relation between attitude and behavioral costs (see, e.g., Kaiser et al., 2020; Taube et al., 2018; Taube & Vetter, 2019), we found psychometric support for the notion that environmental attitude offsets behavioral costs, presumably the single most unique feature of the Campbell paradigm (see, e.g., Kaiser, *in press*; Kaiser et al., 2010). Specifically, we found that cost estimates mirrored increasing actual waiting times (see Figure 3), and estimates of people's environmental attitude corresponded not only with increasing cost estimates but also with increasing actual waiting times (see Figure 2).

According to the Campbell paradigm, people disclose their personal levels of an attitude in the extent to which they actually engage in increasingly more and more challenging behaviors (see Kaiser et al., 2010; Kaiser & Wilson, 2019). In prior research, Campbell-paradigm-based measures were found to converge with conventional attitude scales (see Kaiser et al., 2007, 2013) and with other Campbell-paradigm-based measures even without a single overlapping item (see Kaiser et al., 2018). In this study, in contrast to this prior

research, and for the first time, we were able to corroborate the validity of the cost estimates as a measure of objective costs (i.e., actual waiting times) in otherwise identical pro-environmental behaviors (see Table 2). By doing so, we tested the interpretation of cost estimates and of their relevance for the measurement of environmental attitude in accordance with the Campbell paradigm. In other words, with our research, we were able to connect the cost and attitude estimates, which were established in the measurement process, to actual waiting time costs, which were experimentally manipulated outside the measurement process.

As the cost estimates of the strict and lenient views of pro-environmental behavior were nearly perfectly linearly related to each other (see Figure 3), we have reason to believe that the dichotomization of the PEBT items affected the cost estimates in the expected systematic manner, making one set of items more lenient than the other (see Table 2A in the Appendix). For our test, it was, thus, inconsequential how we ultimately defined pro-environmental engagement (i.e., leniently or strictly).

As is true for any empirical research, our investigation comes with shortcomings. First, our mainly student sample was selective at least in terms of age, education, and income and is, thus, certainly not representative of the population at large. Second, as the data were collected as part of a larger study (see Lange & Dewitte, 2020), we cannot entirely rule out the possibility that participants' responses to the GEB and the PEBT items were affected by other parts of the study. As both these limitations call into question the generalizability of our findings to the population at large, future replications will have to explore whether our findings hold up with other samples from Belgium and other countries.

Third, the environmental consequences of the energy consumed by the 12 USB-powered lights and their associated CO₂ emissions were, in absolute terms, rather small. Yet they were sufficient for inducing our sample to choose the pro-environmental options in the anticipated manner (see Figure 2). More importantly, though, the environmental consequences were not what mattered for this research (for examples with more consequential

pro-environmental behaviors, see, e.g., Kaiser et al., 2020; Taube et al., 2018; Taube & Vetter, 2019). What mattered were the actual waiting times determined and manipulated outside the measurement process. These waiting times were the essential assets of the PEBT behaviors that allowed us to test the measurement theory (as represented in the Campbell paradigm) that we employed to measure environmental attitude (see, e.g., Kaiser et al., 2010, 2013).

5. Conclusion

How can we know the psychological essence of our measures? According to Cronbach and Meehl (1955), we have to explore the nomological network—the measurement theory—in which the presumed latent attribute is foreseen to operate. Here, the latent attribute was an attitude, and one of its possible measurement theories was, thus, the Campbell paradigm (e.g., Kaiser et al., 2010; Kaiser & Wilson, 2019). Originally introduced as a self-report measure of people's pro-environmental behavior, the GEB scale (e.g., Kaiser, 1998; Kaiser & Wilson, 2000, 2004) was later reconceptualized as an environmental attitude measure within the Campbell paradigm (e.g., Kaiser et al., 2007, 2010, 2013, 2018). In an endeavor to establish the GEB as a measure of environmental attitude, we believe we have come full circle after substantiating that the cost estimates validly reflect objective costs. By testing what Wilson (2005) called the internal structure of the measurement theory behind environmental attitude, the Campbell paradigm, we gained confidence in the proposed understanding of the GEB as a measure of environmental attitude.

Supplementary data

Supplementary data (Tables 1S-4S) related to this article can be found at <https://...>

References

- Adams, R. J., & Khoo, S.-T. (1996). *QUEST: The interactive test analysis system*. Acer.
- Arnold, O., & Kaiser, F. G. (2018). Understanding the foot-in-the-door effect as a pseudo effect from the perspective of the Campbell paradigm. *International Journal of Psychology*, 53, 157-165. doi: 10.1002/ijop.12289
- American Psychological Association — APA (2017). *Ethical principles of psychologists and code of conduct*. Author.
- Bamberg, S., & Möser, G. (2007). Twenty years after Hines, Hungerford, and Tomera: A new meta-analysis of psycho-social determinants of pro-environmental behavior. *Journal of Environmental Psychology*, 27, 14-25. doi: 10.1016/j.jenvp.2006.12.002
- Bickman, L. (1972). Environmental attitudes and actions. *The Journal of Social Psychology*, 87, 323-324. doi: 10.1080/00224545.1972.9922533
- Campbell, D. T. (1963). Social attitudes and other acquired behavioral dispositions. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 6, pp. 94-172). McGraw-Hill. doi: 10.1037/10590-003
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302. doi: 10.1037/h0040957
- Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, 60, 170-180. doi: 10.1037/0003-066X.60.2.170
- DeCoster, J., Iselin, A.-M. R., & Gallucci, M. (2009). A conceptual and empirical examination of justifications for dichotomization. *Psychological Methods*, 14, 349-366. doi: 10.1037/a0016956
- De Houwer, J., Gawronski, B., & Barnes-Holmes, D. (2013). A functional-cognitive framework for attitude research. *European Review of Social Psychology*, 24, 252-287. doi: 10.1080/10463283.2014.892320

- Diekmann, A., & Preisendörfer, P. (1998). Environmental behavior: Discrepancies between aspirations and reality. *Rationality and Society*, 10, 79–102. doi: 10.1177/104346398010001004
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Erlbaum.
- Gifford, R. (2014). Environmental psychology matters. *Annual Review of Psychology*, 65, 541–579. doi: 10.1146/annurev-psych-010213-115048
- Hastie, R., & Dawes, R. M. (2010). *Rational choice in an uncertain world: The psychology of judgment and decision making* (2nd ed.). Sage.
- Kaiser, F. G. (1998). A general measure of ecological behavior. *Journal of Applied Social Psychology*, 28, 395-422. doi: 10.1111/j.1559-1816.1998.tb01712.x
- Kaiser, F. G. (in press). Climate change mitigation within the Campbell paradigm: Doing the right thing for a reason and against all odds. *Current Opinion in Behavioral Sciences*.
- Kaiser, F. G., & Byrka, K. (2015). The Campbell paradigm as a conceptual alternative to the expectation of hypocrisy in contemporary attitude research. *The Journal of Social Psychology*, 155, 12-29. doi: 10.1080/00224545.2014.959884
- Kaiser, F. G., Byrka, K., & Hartig, T. (2010). Reviving Campbell's paradigm for attitude research. *Personality and Social Psychology Review*, 14, 351-367. doi: 10.1177/1088868310366452
- Kaiser, F. G., Hartig, T., Brügger, A., & Duvier, C. (2013). Environmental protection and nature as distinct attitudinal objects: An application of the Campbell paradigm. *Environment and Behavior*, 45, 369-398. doi: 10.1177/0013916511422444
- Kaiser, F. G., Henn, L., & Marschke, B. (2020). Financial rewards for long-term environmental protection. *Journal of Environmental Psychology*, 68, 101411. doi: 10.1016/j.jenvp.2020.101411
- 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. doi: 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. doi: 10.1016/j.jenvp.2007.06.004
- Kaiser, F. G., & Wilson, M. (2000). Assessing people's general ecological behavior: A cross-cultural measure. *Journal of Applied Social Psychology*, 30, 952-978. doi: 10.1111/j.1559-1816.2000.tb02505.x
- Kaiser, F. G., & Wilson, M. (2004). Goal-directed conservation behavior: The specific composition of a general performance. *Personality and Individual Differences*, 36, 1531-1544. doi: 10.1016/j.paid.2003.06.003
- Kaiser, F. G., & Wilson, M. (2019). The Campbell paradigm as a behavior-predictive reinterpretation of the classical tripartite model of attitudes. *European Psychologist*, 24, 359-374. doi: 10.1027/1016-9040/a000364
- Klößner, C. A., & Blöbaum, A. (2010). A comprehensive action determination model: Toward a broader understanding of ecological behavior using the example of travel mode choice. *Journal of Environmental Psychology*, 30, 574-586. doi: 10.1016/j.jenvp.2010.03.001
- 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. doi: 10.1080/13504620220145401
- Kormos, Ch, & Gifford, R. (2014). The validity of self-report measures of proenvironmental behavior: A meta-analytic review. *Journal of Environmental Psychology*, 40, 359-371. doi: 10.1016/j.jenvp.2014.09.003

- Lange, F., & Dewitte, S. (2020). Positive affect and pro-environmental behavior: A preregistered experiment. *Journal of Economic Psychology*, 80, 102291. doi: 10.1016/j.joep.2020.102291
- Lange, F., & Dewitte, S. (2019)a. Cognitive flexibility and pro-environmental behaviour: A multimethod approach. *European Journal of Personality*, 33, 488-505. doi: 10.1002/per.2204
- Lange, F., & Dewitte, S. (2019)b. Measuring pro-environmental behavior: Review and recommendations. *Journal of Environmental Psychology*, 63, 92-100. doi: 10.1016/j.jenvp.2019.04.009
- 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. doi: 10.1016/j.jenvp.2018.02.007
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7, 19-40. doi: 10.1037/1082-989X.7.1.19
- 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. doi: 10.1037/h0034936
- Maloney, M. P., Ward, M. P., & Braucht, G. N. (1975). Psychology in action: A revised scale for the measurement of ecological attitudes and knowledge. *American Psychologist*, 30, 787-790. doi: 10.1037/h0084394
- Matell, M. S., & Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity. *Educational and Psychological Measurement*, 31, 657-674. doi: 10.1177/001316447103100307

- Milfont, T. L., & Duckitt, J. (2004). The structure of environmental attitudes: A first and second-order confirmatory factor analysis. *Journal of Environmental Psychology, 24*, 289-303. doi: 10.1016/j.jenvp.2004.09.001
- Peabody, D. (1962). Two components in bipolar scales: Direction and extremeness. *Psychological Review, 69*, 65-73. doi: 10.1037/h0039737
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. University of Chicago Press (Original work published 1960).
- Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1–14). Yale University Press.
- Schultz, P. W. (2001). The structure of environmental concern: Concern for self, other people, and the biosphere. *Journal of Environmental Psychology, 21*, 327-339. doi: 10.1006/jevp.2001.0227
- Schultz, P. W., & Oskamp, S. (1996). Effort as a moderator of the attitude–behavior relationship: General environmental concern and recycling. *Social Psychology Quarterly, 59*, 375–383. doi: 10.2307/2787078
- Stern, P. C. (2000)a. Psychology and the science of human–environment interactions. *American Psychologist, 55*, 523–530. doi: 10.1037//0003-066X.55.5.523
- Stern, P.C. (2000)b. Toward a coherent theory of environmentally significant behavior. *Journal of Social Issues, 56*, 407–424. doi: 10.1111/0022-4537.00175
- 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. doi: 10.1016/j.trf.2018.05.006

- Taube, O., & Vetter, M. (2019). How green defaults promote environmentally friendly decisions: Attitude-conditional default acceptance but attitude-unconditional effects on actual choices. *Journal of Applied Social Psychology, 49*, 721-732. doi: 10.1111/jasp.12629
- Whitely, S. E. (1977). Models, meanings and misunderstandings: Some issues in applying Rasch's theory. *Journal of Educational Measurement, 14*, 227-235. doi: 10.1111/j.1745-3984.1977.tb00040.x
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Erlbaum.
- Wright, B. D. (1977). Misunderstanding the Rasch model. *Journal of Educational Measurement, 14*, 219-225. Retrieved from www.jstor.com/stable/1434313
- Wright, B. D., Linacre, J. M., Gustafson, J.-E., & Martin-Löf, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), 370. Retrieved from www.rasch.org/rmt/rmt83b.htm
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. MESA.
- Zhu, X., & Lu, C. (2017). Re-evaluation of the New Ecological Paradigm scale using item response theory. *Journal of Environmental Psychology, 54*, 79-90. doi: 10.1016/j.jenvp.2017.10.005

Endnotes

- ¹ Although Cronbach and Meehl's logic is typically applied to substantive theory, we are unaware of any reason why measurement theories cannot be equally tested when they involve observables.
- ² As a derivative from behavior, Campbellian attitude measures are akin to what economists would presumably call *revealed preference* measures (see Hastie & Dawes, 2010). Unlike revealed preferences in economy, people's particular level of an attitude cannot be reliably revealed by a single behavior within the Campbell paradigm.
- ³ The model in Figure 1 does not imply that a specific behavior (e.g., commuting by bike) is exclusively aimed at environmental protection. Commuting by bike can just as well be aimed at protecting one's health or at saving money, to name two other possibilities. We call it a pro-environmental behavior, though, when bike commuting is aimed at environmental protection and thus reflects a person's environmental attitude (see the intent-oriented definition of behavior by Stern, 2000b). If bike commuting was aimed at health protection, we would call it a health behavior because health attitude and, thus, protecting one's health was what triggered it (Kaiser, in press).

The first step in identifying a particular attitude (e.g., environmental attitude) as the origin of behavior requires researchers to take stock of an array of behaviors people engage in to protect the environment (see Kaiser, in press). This is why measures of environmental attitude within the Campbell paradigm typically consist of self-reports of past more or less taxing environmentally protective behaviors (e.g., the General Ecological Behavior scale; e.g., Kaiser, 1998; Kaiser & Wilson, 2000, 2004). In a second step, person scores tallied with this measure must be empirically tested in the nomological network (see Cronbach & Meehl, 1955) in which environmental attitude is theoretically anticipated to operate by, for example, accounting for manifest

- environmentally protective activities not employed in the measurement of the attitude in question (see, e.g., Kaiser et al., 2020; Taube et al., 2018; Taube & Vetter, 2019).
- 4 All Rasch model tests and scale calibrations were performed with QUEST 2.1 (Adams & Khoo, 1996).
 - 5 An inspection of the convergent validity of two measures of environmental attitude grounded in different subsets of GEB items (one grounded in the 31 positively formulated items and one grounded in the 19 negatively formulated items) revealed a measurement-error-corrected correlation of $r_{\text{corr}} = .63$ (uncorrected: $r = .41$) between the two scales. This correlation approached the average person separation reliability of the two scales of $rel = .65$. Accordingly, the correlation speaks of two fairly invariant scales.
 - 6 Of the 32 GEB items with 5-point frequency response options, 19 were arranged in a disorderly fashion. In these items, the engagement frequency order (from never to very often) did not correspond with increasing cost estimates. As a consequence, the variance of the attitude estimates was cut into a fourth (from .64 to .17). At the same time, the separation reliability of the persons did not improve much (from .76 to .80) despite the substantial extension in the available information on which the attitude estimates were grounded (146 instead of 50 cost estimates; see Tables 1S-4S in the Supplementary data).
 - 7 In a subsequent test—using the partial credit Rasch model—we demonstrated that our dichotomization practice was once again successful at preventing invalid response patterns. Specifically, we found that the cost estimates of the engagement frequencies (1 to 4) in the six PEBT items were by and large inversely related to the actual cumulative waiting times. This meant that longer waiting times were counterintuitively related to fewer behavioral costs and, as such, waiting only once in four trials was more (instead of less) demanding than waiting four times in four trials (see Table 3S in the Supplementary data).

Appendix

Table 1A

The 50 Self-Reports of the General Ecological Behavior Scale Ordered by their Estimated Costs

	BEHAVIORAL SELF-REPORTS	δ	MS_w	MS_u	t_w	t_u
1.	I buy domestically grown wooden furniture.	3.73 (.60)	1.00	0.63	0.2	-0.3
2.	I contribute financially to environmental organizations.	3.32 (.37)	1.00	0.89	0.1	-0.1
3.	I drive on freeways at speeds under 100 kph (= 62.5 mph).	2.76 (.38)	1.02	1.06	0.2	0.3
4.	I am a member of an environmental organization.	2.51 (.26)	0.91	0.78	-0.4	-0.7
5.	I boycott companies with an unecological background.	2.44 (.27)	0.89	0.60	-0.5	-1.5
6.	I am a vegetarian.	2.10 (.23)	0.94	0.84	-0.4	-0.6
7.	I am a member of a carpool.	1.83 (.24)	1.12	1.13	0.8	0.6
8.	I buy meat and produce with eco-labels.	1.83 (.22)	0.93	0.85	-0.4	-0.7
9.	I buy products in refillable packages.	1.62 (.20)	0.96	1.00	-0.3	0.1
10.	I own solar panels.	1.47 (.21)	1.12	1.19	1.0	1.0
11.	<i>At red traffic lights, I keep the engine running.</i>	1.42 (.23)	0.98	0.94	-0.1	-0.2
12.	<i>I buy convenience foods.</i>	1.24 (.18)	1.08	1.12	0.8	0.8
13.	I have a contract for renewable energy with my energy provider.	1.17 (.25)	0.91	0.85	-0.7	-0.7
14.	<i>For longer journeys (more than 6 hours of travel time by car), I take an airplane.</i>	0.95 (.17)	1.13	1.33	1.7	2.3
15.	I buy beverages and other liquids in returnable bottles.	0.84 (.17)	0.99	0.99	-0.2	-0.1
16.	I have pointed out unecological behavior to someone.	0.76 (.17)	0.93	0.93	-1.0	-0.6

continued

BEHAVIORAL SELF-REPORTS		δ	MS_w	MS_u	t_w	t_u
17.	<i>I use a clothes dryer.</i>	0.66 (.17)	1.07	1.26	1.1	2.1
18.	I talk with friends about environmental pollution, climate change, and/or energy consumption.	0.57 (.16)	0.90	0.85	-1.8	-1.4
19.	I read about environmental issues.	0.52 (.16)	0.86	0.82	-2.5	-1.7
20.	I refrain from owning a car.	0.51 (.20)	1.12	1.18	1.7	1.3
21.	<i>I keep the engine running while waiting in front of a railroad crossing or in a traffic jam.</i>	0.51 (.20)	0.90	0.86	-1.4	-1.1
22.	I own a fuel-efficient automobile (less than 6 liters per 100 kilometers).	0.48 (.24)	0.95	0.91	-0.6	-0.5
23.	I have looked into the pros and cons of having a private source of solar power.	0.47 (.18)	0.88	0.87	-1.9	-1.1
24.	<i>I use fabric softener with my laundry.</i>	0.44 (.18)	1.12	1.12	1.8	1.0
25.	<i>If I am offered a plastic bag in a store, I take it.</i>	0.14 (.16)	0.90	0.88	-1.9	-1.2
26.	<i>I drive to where I want to start my hikes.</i>	-0.07 (.21)	1.25	1.43	3.6	2.9
27.	<i>I buy beverages in cans.</i>	-0.19 (.16)	1.02	1.02	0.4	0.2
28.	<i>In the winter, I keep the heat on so that I do not have to wear a sweater.</i>	-0.25 (.16)	0.98	0.95	-0.4	-0.5
29.	I collect and recycle used paper.	-0.37 (.16)	1.00	1.00	0.0	0.1
30.	I drive in such a way as to keep my fuel consumption as low as possible.	-0.40 (.20)	0.86	0.80	-2.2	-1.6
31.	I own an energy efficient dishwasher (efficiency class A+ or better).	-0.49 (.23)	1.02	1.02	0.3	0.2
32.	I buy seasonal produce.	-0.55 (.16)	0.99	1.03	-0.1	0.3
33.	<i>I drive my car in or into the city.</i>	-0.60 (.22)	1.09	1.19	1.1	1.2
34.	<i>I buy bleached or colored toilet paper.</i>	-0.79 (.18)	1.04	1.16	0.6	1.2

continued

	BEHAVIORAL SELF-REPORTS	δ	MS_w	MS_u	t_w	t_u
35.	In nearby areas (around 30 kilometers; around 20 miles), I use public transportation or ride a bike.	-0.87 (.17)	1.07	1.06	1.0	0.5
36.	In winter, I turn down the heat when I leave my apartment/house for more than 4 hours.	-0.91 (.17)	1.01	1.03	0.1	0.3
37.	I wash dirty clothes without prewashing.	-0.92 (.20)	1.10	1.22	1.2	1.4
38.	<i>I use a chemical air freshener in my bathroom.</i>	-1.00 (.18)	1.03	0.97	0.4	-0.2
39.	I bring empty bottles to a recycling bin.	-1.18 (.18)	0.98	0.87	-0.2	-0.8
40.	<i>I kill insects with a chemical insecticide.</i>	-1.19 (.19)	0.99	0.95	-0.1	-0.3
41.	<i>In hotels, I have the towels changed daily.</i>	-1.19 (.19)	0.92	0.93	-0.8	-0.4
42.	<i>I use an oven cleaning spray to clean my oven.</i>	-1.28 (.21)	1.00	1.07	0.0	0.5
43.	<i>In the winter, I air rooms while keeping the heat on and leaving the windows open, simultaneously.</i>	-1.40 (.19)	0.98	0.91	-0.1	-0.5
44.	<i>I put dead batteries in the garbage.</i>	-1.97 (.23)	1.00	1.06	0.0	0.3
45.	I shower (rather than take a bath).	-2.18 (.24)	1.01	1.69	0.1	2.3
46.	<i>After meals, I dispose of leftovers in the toilet.</i>	-2.64 (.28)	0.92	.78	-0.3	-0.6
47.	I wait until I have a full load before doing my laundry.	-2.95 (.35)	1.00	0.92	0.1	0.0
48.	After a picnic, I leave the place as clean as it was originally.	-3.01 (.33)	0.99	0.91	0.1	-0.1
49.	I ride a bicycle, walk, or take public transportation to work or school.	-3.94 (.51)	0.98	3.03	0.1	2.2
50.	I reuse my shopping bags.	-3.97 (.51)	1.04	1.26	0.2	0.6

Note. Self-reports in italics indicate negatively formulated behaviors; they are recoded and should be read as "I refrain from" Cost estimates (δ) are expressed in logits. Mean square (MS) and t values—unweighted ($_u$) and weighted ($_w$) by the item variance—reflect the relative discrepancy between the model predictions and the observed data and were, thus, used to assess item fit (see, e.g., Wright & Masters, 1982). **Bold figures** highlight either statistically significant t values ($p < .05$) or MS_w values that did not fall within an acceptable range of fit even for test items ($0.80 \leq MS \leq 1.20$) and MS_u values that did not fall within an acceptable

range for typical survey items ($0.60 \leq MS \leq 1.40$; see Wright et al., 1994). A value of $MS > 1.20$ corresponds to an excess of more than 20% variation in the model's prediction compared with what was in the data.

Table 2A

Cost Estimates of the Six PEBT Behaviors (i.e., Bicycle Choice) with Different Waiting Times when Leniently or Strictly Defined

WAITING TIME FOR BICYCLE CHOICES:	δ_{lenient}	Upper bound	Lower bound	δ_{strict}
60 s	0.60	0.91	1.67	2.11
30 s	-0.39	-0.08	0.93	1.26
20 s	-1.12	-0.75	0.42	0.73
15 s	-2.11	-1.62	-0.22	0.09
10 s	-3.49	-2.59	-1.27	-0.92
5 s	-5.10	-3.14	-2.53	-2.04

Note. All numerical information has been converted into the metric of the lenient instrument to make the numbers directly comparable. Pro-environmental behaviors (i.e., bicycle choices) are ordered along decreasing waiting times (in s). Estimates of behavioral costs are expressed in logits for behaviors with lenient (δ_{lenient}) and strict (δ_{strict}) definitions of pro-environmental behavior. The upper bound stands for the upper end of the 95% confidence interval of cost estimates with a lenient interpretation of pro-environmental behavior. The lower bound stands for the lower end of the 95% confidence interval of cost estimates with a strict interpretation of pro-environmental behavior. All comparisons represent significantly distinct cost estimates at $p \leq .01$ (i.e., all upper bound figures are smaller than their counterpart lower bound figures).