MEASURING BIOSPHERIC VALUES:
ITALIAN ADAPTATION OF THE GREEN SCALE

MATTEO ROBBA
PAOLA IANNELLO
ANGELA SORGENTE
CATHOLIC UNIVERSITY OF MILANO, ITALY

Climate change is requiring people to shift toward more sustainable lifestyles. Although the contribution of behavioral sciences in sustainability domains is growing, validated instruments to measure biospheric values are still few, especially in Italian. The present study aims to test the psychometric properties of the GREEN scale in Italian, by adopting the best practices suggested by the contemporary view of validity. Data was retrieved from a representative sample of 1,002 Italian individuals, split into two subgroups. Both exploratory and confirmatory factor analyses supported the original factorial structure of the scale and the measurement invariance showed that this structure is equivalent across genders. Furthermore, convergent, discriminant, and criterion-related validity evidence demonstrated that the measure is associated with universalism, connectedness to nature, and daily pro-environmental behaviors, respectively. Collected evidence suggests that the GREEN scale is a valid and reliable measure for biospheric values.

Keywords: Biospheric values; Pro-environmental behaviors; Connectedness to nature; Sustainability; Validation.

Correspondence concerning this article should be addressed to Matteo Robba, Department of Psychology, Catholic University, Largo Gemelli 1, 20123 Milano (MI), Italy. Email: matteopaolo.robba@unicatt.it

Climate change is one of the biggest threats that mankind is called to deal with from now to the next few decades, requiring people to shift their habits by pursuing more environmentally-friendly behaviors and consumption. In the last decade, the psychological literature on sustainability has grown considerably. Specifically, a relevant research field concerns the investigation of psychological characteristics associated with pro-environmental behaviors. Various researchers highlighted the relevant role played by personal values in adopting environmentally-friendly behaviors and sustainable consumption habits. According to Schwartz (1992), values function as guiding principles in individuals' lives, influencing their actions. Furthermore, values are abstract and general constructs that apply to a wide range of contexts, because they transcend specific domains and situations. Concerning environmental domains, scholars generally talk about biospheric (or environmental) values, namely self-transcendent values reflecting the concern for caring for and preserving the natural environment (De Groot & Steg, 2010).

Previous studies reported a strong association between them and environmental protection behaviors (De Groot & Steg, 2009; Katz-Gerro et al., 2017; Ruepert et al., 2017; Wang et al., 2021). For instance, Bouman et al. (2020) showed in a cross-country study that biospheric values are related to both worry about climate change and energy-saving behaviors. Likewise, consumers with stronger biospheric values are more prone to purchasing organic products (van Doorn & Verhoef, 2015). Values can also influence individuals' attitudes. Indeed, a strong impact of values on environmental attitudes was reported by various studies (e.g., Schultz & Zelezny, 1999; Steg et al., 2014). For instance, in a cross-cultural comparison, Schultz et al. (2005)

found a significant association between values and environmental concern, that is, a general attitude referring to a personal evaluation of environmental issues (De Groot & Thøgersen, 2018). Biospheric values are also related to a willingness to sacrifice personal interests for the environment (Rahman & Reynolds, 2016) and to think about the environmental consequences of one's actions (Perlaviciute & Steg, 2015).

PSYCHOMETRIC INSTRUMENTS MEASURING BIOSPHERIC VALUES

Despite the topic relevance, only a few psychometric scales assess values in environmental domains. The most widely used instruments were created relying on Schwartz's value theory. Specifically, according to this theoretical framework (Schwartz, 1992, 1994), values can be distinguished between self-enhancement and self-transcendence. While the former are more focused on the concerns for personal interests, the latter refer to stronger interests for the collectivity. These two types of values are differently associated with pro-environmental behaviors. Indeed, while self-transcendence values positively affect environmentally-friendly behaviors, self-enhancement values are typically negatively related to them (Bouman et al., 2018).

Steg et al. (2014) proposed the Environmental-SVS (E-SVS), which has been adapted from the Schwartz Value Survey (SVS; Schwartz, 1994). This tool consists of 16 items measuring four values. Indeed, as suggested by the authors, four specific values are relevant in predicting pro-environmental behaviors. Among self-transcendent values, there are biospheric and altruistic (i.e., concerns for the welfare and fair treatment of other individuals) values. Among self-enhancement values, there are egoistic and hedonic values. The former refers to the maximization of personal outcomes, while hedonic values focus on the search for pleasure and positive feelings (De Groot & Thøgersen, 2018). The SVS received various critiques over time, and the same concerns could be applied to the E-SVS as well (Bouman et al., 2018). Specifically, the items are developed in a way that might enhance the self-presentation bias in the answers. Furthermore, respondents may find it difficult to answer the items of the scale, thus affecting the validity and reliability of the measure itself. To overcome the limits of the SVS, Schwartz et al. (2012) proposed the Portrait Values Questionnaire (PVQ), which is currently considered the gold standard for values assessment. Among the 10 different motivational behaviors, the PVQ measures universalism, in which altruistic and biospheric values are embedded and assessed with three items each. Recently, Bouman et al. (2018) adapted the PVQ to specifically measure values in environmental domains by developing the Environmental Portrait Values Questionnaire (E-PVQ). This instrument assesses four types of values: biospheric, altruistic, hedonic, and egoistic. To date, neither the E-SVS nor the E-PVQ have been adapted yet to the Italian context.

However, studies report that, among these values, the biospheric ones are the strongest predictors of pro-environmental behaviors (e.g., Katz-Gerro et al., 2017), thus suggesting the possibility to specifically focus on them. Furthermore, in some situations, altruistic and biospheric values could be conflicting, for instance when primed at the same time (De Groot & Thøgersen, 2018; van Doorn & Verhoef, 2015). As suggested by De Groot and Steg (2007), though related, biospheric and altruistic are distinct types of self-transcendence values and, in some contexts, they may contribute uniquely to explaining human behavior. Therefore, it may be useful to consider biospheric and altruistic values independently, given that they have a different focus.

A specific measure for biospheric values is the GREEN scale (Haws et al., 2014), a brief tool reflecting the extent to which individuals pay attention to the conservation of natural resources in their daily behavior. The measure is composed of six items, developed on a 7-point Likert scale, and it was originally

TPM Vol. 31, No. 3, September 2024
345-361
© 2024 Cises

Robba, M., Iannello, P.,
& Sorgente, A.
Italian adaptation of the GREEN scale

proposed in English. This scale has been used in various domains. For instance, it positively predicted attitudes toward environmentally-friendly products (Pegan et al., 2023), like cultured meat (Dupont et al., 2022) and bio-based apparel (Stahl et al., 2021). Likewise, the GREEN scale was found to predict sustainable consumption behaviors (e.g., Dikici et al., 2022; Paço et al., 2019).

This instrument has been used in various countries, such as the United States (Bailey et al., 2018), Germany (Klein et al., 2020; Macht et al., 2023), the United Kingdom (Paço et al., 2019; Ribeiro et al., 2023), Austria (Essiz & Mandrik, 2021), Poland (Bartoszczuk et al., 2022), and Turkey (Dikici et al., 2022). The measure was adopted for research purposes within the Italian context as well (e.g., Bonera et al., 2020; Mazzocchi et al., 2021; Risitano et al., 2023), though without testing its psychometric properties. This data further suggests the usefulness of an Italian adaptation of the GREEN scale. Indeed, to date, there is no validation of the measure, despite its good psychometric properties and its brevity allow researchers to use it in different settings and easily integrate it in longer surveys. An Italian adaptation of the scale could be useful to both enhance our theoretical understanding of how values shape human behavior in environmental domains (e.g., cross-cultural comparisons) and develop interventions to promote pro-environmental behaviors. Likewise, it could also be useful in marketing and sustainable consumption research.

STUDY PURPOSE

The present study aimed to collect validity evidence for the Italian version of the GREEN scale (Haws et al., 2014). The validation process was carried out following the contemporary view of validity (Hubley & Zumbo, 2011) guidelines. Hence, different kinds of validity evidence were collected to evaluate the construct validity of the instrument.

THEORETICAL FRAMEWORK

The conception of validity has changed multiple times during the last century and the different conceptions in vogue nowadays share some common assumptions that are identified by the expression "contemporary view of validity." According to Zumbo (2005, 2006; Sorgente & Zumbo, in press) the three main differences between the traditional view(s) of validity and the contemporary view(s) of validity are the following. First, validity is no longer a property of a test. According to the traditional view of validity, once the test developer has demonstrated that a test measures what it claims to measure, the test is valid per se, and it can be adopted in future studies with the inherent assurance that the instrument is adequate to assess the construct of interest. However, according to the contemporary view of validity, "trying to define validity as a property of tests quickly leads to absurdities since the validity of a test can vary from population to population" (Wiliam, 2014, p. 29). In other words, a researcher has to consider whether there is sufficient evidence in the research literature to "support the appropriateness, meaningfulness, and usefulness of the specific inferences made from scores about individuals from a given sample and in a given context" (Zumbo, 2006, p. 48), and, if not, that evidence needs to be provided. This implies that validity is contextualized; we cannot validate a test per se, because the validity of a test may change from one context to another (e.g., different countries, samples, and age groups).

This contextualized validity implies a second difference between the traditional and contemporary views of validity. Whereas according to the traditional view, test developers are the only ones who must evaluate whether an instrument is valid or not, the contemporary view also calls for test users (i.e., researchers,

clinicians, and practitioners adopting the psychometric instrument), together with test developers, to prove the validity of the scores obtained from a test in a specific sample or context.

Finally, the contemporary view proposed a unified conception of validity. According to the traditional view of validity, different kinds of validity exist (e.g., content, predictive, concurrent, and construct validity), and each of these "validities" is treated as sufficient evidence on its own. The contemporary view of validity (Messick, 1989, 1995) suggests instead that test developers and users should specifically evaluate the *construct* validity of their scores. Because "construct validity cannot generally be expressed in the form of a single simple coefficient" (Cronbach & Meehl, 1955, p. 201), the validation process requires collecting multiple sources of validity evidence to prove that the test scores measure the construct they prompt to measure (i.e., construct validity). In particular, one may formulate a list of hypotheses to empirically demonstrate the theory (e.g., the construct should have X dimensions, be associated with variable Y, predict variable Z, etc.). Testing these hypotheses and finding evidence to support them means collecting evidence of construct validity. This description focuses on different kinds of evidence (e.g., convergent evidence) rather than diverse kinds of validity (e.g., convergent validity).

Hubley and Zumbo (2011) have proposed a systematic validation practice (see Figure 1 in Hubley & Zumbo, 2011) in agreement with this unified and contemporary view of validity. They suggest that, to prove construct validity, different hypotheses should be formulated based on the literature available about the construct the instrument prompts to measure and evidence should be collected to confirm these hypotheses:

1. *Content evidence* consists in developing (or translating) items that the researcher hypothesizes, based on the literature, should represent a specific construct and then asking experts to evaluate the appropriateness of those items.

2. *Score structure evidence* consists in hypothesizing the number of dimensions an instrument should have and confirming this factorial structure. Furthermore, the researcher should prove that each dimension of this instrument measures the construct in a precise and consistent way (reliability evidence).

3. *Generalizability evidence* consists in hypothesizing that the instrument measures the same construct across different subgroups of participants (e.g., gender).

4. *Known groups evidence* consists in comparing the mean level of a test score across groups. Researchers should demonstrate that test scores mean levels are different across these groups in case this difference is expected by the literature.

5. *Convergent and discriminant evidence* consists in proving that the test scores of the newly developed instruments are strongly correlated with instruments that, according to the literature, measure the same construct (convergent evidence) as well as that test scores are poorly correlated with instruments that measure a different construct (discriminant evidence).

6. *Criterion-related evidence* consists in demonstrating that the test scores of the newly developed instrument can predict a criterion variable that, according to the literature, should depend on the construct the new instrument aims to measure.

After testing all these hypotheses, the researchers should evaluate whether "all of the accumulated evidence supports the intended interpretation of test scores for the proposed purpose" (Hubley & Zumbo, 2011, p. 220). In other words, if the various kinds of evidence collected confirm the hypotheses suggested by the literature, the researcher can conclude that the instrument measures the constructs it claims to measure (construct validity), at least in the context and sample it was tested on. If a test user aims to use the same instrument in a different context (e.g., country) or sample (e.g., different age range), he/she should prove again the construct validity of the scores obtained from this instrument (i.e., collect new evidence) before interpreting its scores.

TPM Vol. 31, No. 3, September 2024
345-361
© 2024 Cises

Robba, M., Iannello, P.,
& Sorgente, A.
Italian adaptation of the GREEN scale

This paragraph summarizes how the contemporary view of validity has strongly changed the theoretical assumptions (e.g., validity is contextualized) as well as the practices about validation (e.g., many kinds of evidence need to be collected). It is important to point out that the last decades have also offered psychometricians many new and advanced statistical techniques, which can be adopted in validation practice. According to Zumbo (2005), a class of techniques is particularly central to the contemporary validation process: structural equation models (SEM). The traditional procedures widely utilized the Pearson correlation between *observed* measures. For example, the correlation between two instruments (e.g., the instrument to validate and a gold standard or a criterion) or between two scores from the same instrument (e.g., test-retest correlation). For the contemporary view of validity, these correlational analyses are useful, but not sufficient (Zumbo, 2005). First, the correlations among observed measures can be substituted with correlations among *latent* variables (correlations created using SEM), excluding the measurement error from the correlation estimate. Furthermore, analyses other than correlations are applied to validate test scores, and all of these can be performed using SEM: for example, confirmatory factor analysis, measurement invariance, and composite reliability.

METHODS

Procedure

To validate the Italian version of the GREEN scale, we first took care of the item translation and back translation following the suggested guidelines (Geisinger, 1994; van Widenfelt et al., 2005). We then collected content validity evidence by asking experts to evaluate the comprehensibility and representativeness of translated items. Specifically, the translated items were revised by a team of experts in psychometrics and/or psychology of sustainability to evaluate item quality in terms of comprehensibility (i.e., how easy and immediate it is to understand the proposed statement) and representativeness (i.e., how important and coherent the proposed statement is in light of the research objectives previously presented and for each area of investigation). All the items were considered fully adequate, and no further modifications were made.

Once the translated items were approved, we administered them, together with other measurement scales, to a representative sample of 1,002 Italians. This sample was split into two subsamples. The first was adopted to perform an exploratory factor analysis (EFA) and items analysis of the translated items. Once we verified that all items were good and the mono-dimensional structure of the scale was meaningful in the Italian context too, we conducted the validation of the GREEN scale adopting the second subsample. As reported above, the construct validity of a psychometric scale can only be assessed by collecting different kinds of evidence. Hence, a confirmatory factor analysis (CFA) was first performed to collect score structure validity evidence. Omega coefficient and average variance extracted (AVE) were estimated to collect reliability evidence. Then, measurement invariance across genders was assessed to test generalizability evidence.

To test known groups evidence, average values were compared between males and females. Gender differences in pro-environmental behaviors are a well-known phenomenon, because women are more likely to play out environmentally-friendly behaviors (Briscoe et al., 2019; Vicente-Molina et al., 2013, 2018; Xia & Li, 2023). Though gender differences in biospheric values are less explored, existing studies show a similar trend, with women reporting higher levels than men. However, it is worth highlighting that these differences are generally marginal (e.g., Milfont & Sibley, 2016). For instance, a recent cross-country comparison performed in Europe (Sargisson et al., 2020) found only a weak association between gender and biospheric values. Indeed, women are significantly more concerned about protecting the environment, but the difference

TPM Vol. 31, No. 3, September 2024
345-361
© 2024 Cises

Robba, M., Iannello, P.,
& Sorgente, A.
Italian adaptation of the GREEN scale

is small. While this evidence is mainly related to pro-environmental *behavior* and not specifically to biospheric *values,* we expected this difference may be reported by the GREEN scale as well.

Subsequently, three structural equation models (SEM) were estimated to obtain convergent, discriminant, and criterion-related validity evidence. For each of the three models, latent variables were considered to remove the measurement error from these validity estimates (Zumbo, 2005). To collect convergent evidence, a correlation between the target scale scores and the scores of a psychometric instrument measuring the same construct should be performed to prove that the two measures are closely related. Consequently, we tested the correlation between the GREEN scale and an instrument (i.e., the UN subscale from the PVQ-RR; Schwartz & Cieciuch, 2022) measuring universalism, namely a personal value reflecting attention to preserving people's welfare and the natural environment (Schwartz et al., 2012). We specifically focused on the measure of universalism-nature (UNN). Because it emphasizes the importance of safeguarding and living in harmony with the natural world, the latent scores of the GREEN scale and the UNN subscale were expected to be strongly associated. To check whether the GREEN scale can discriminate between different constructs (i.e., discriminant evidence), a second correlation model with latent variables was run. In particular, the relationship between biospheric values and connectedness to nature was assessed. This construct refers to how much someone feels a sense of belonging in the natural environment (Lengieza & Swim, 2021). Though related, these two concepts are theoretically distinct. Indeed, differently from biospheric values, which reflect concerns for and importance of nature protection, connectedness to nature has a greater focus on an individual's subjective and emotional experience of oneness with the natural world. Conversely, biospheric values are more abstract and cognitive, functioning as beliefs and guiding principles toward nature preservation (Martin & Czellar, 2017; van der Werff et al., 2013). Accordingly, a significant association between the latent scores of the GREEN scale and the Illustrated Inclusion of Nature in Self scale (IINS; Kleespies et al., 2021) was expected, though with a lower correlation coefficient compared to the convergent one. Finally, the relationship between biospheric values and pro-environmental behaviors was estimated to collect criterion evidence. The literature suggests that values are associated with pro-environmental behaviors and sustainable consumption habits (e.g., De Groot & Steg, 2009). A moderate and positive association between the latent scores of the GREEN scale and the measure for pro-environmental behaviors was thus expected.

Sample

Data was retrieved from a representative sample of 1,002 Italian individuals, who filled out an online survey in June 2023. Participants were required to give written informed consent. The survey completion took approximately 10 minutes. The present study was approved by the Ethical Committee of Catholic University and followed the American Psychological Association (APA) standard ethical guidelines for research. Respondents had to be of legal age (i.e., $\geq$ 18 years old). No other inclusion/exclusion criteria were applied.

To ensure sample representativeness for gender, age, education, and geographical area, a quota sampling procedure was adopted. Females ($n = 500$) accounted for 49.9% of participants. The age of the sample ranged from 18 to 54, with an average of 37.19 years ($SD = 10.94$). Over half of the sample (52%) had a high school diploma, 31.4% of respondents attended university, while the remaining 16.6% had a junior-high

school degree. Finally, as for geographical area, 26.4% of participants lived in the north-west of Italy, 19.4% in the north-east, 22.4% in the center, and 31.8% in the southern regions.

For study purposes, respondents were randomly divided into two groups. On the first subsample an exploratory factor analysis and item analysis ($n = 503$) was performed, while the second one was used to collect validity evidence ($n = 499$). Table 1 summarizes the sociodemographic characteristics of the two subsamples.

TABLE 1
Sociodemographic characteristics of the two subsamples

	Subsample 1 ($n = 503$)	Subsample 2 ($n = 499$)
Gender		
Males	49.7%	50.5%
Females	50.3%	49.5%
Age	$M = 36.69$ ($SD = 11.06$)	$M = 37.68$ ($SD = 10.79$)
Education		
Junior-high school	15.5%	17.6%
High school	51.7%	52.3%
University degree	32.8%	30.1%

Note. M = mean; SD = standard deviation.

Measures

Biospheric values were assessed using the GREEN scale. This instrument, developed by Haws et al. (2014), consists of six items rated on a 7-point Likert scale, with answers ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The measure is composed of a single factor and higher scores indicate greater levels of concern for the environment. The Italian version of the scale is reported in the Appendix.

To measure *nature universalism*, the UNN subscale from the Revised Portrait Value Questionnaire (PVQ-RR; Schwartz & Cieciuch, 2022) was adopted. The subscale is formed by three items ($\omega = .87$), estimating individuals' attention to preserve the natural environment (e.g., "It is important to him/her to protect the natural environment from destruction or pollution"). Respondents are asked to express their agreement on a 7-point Likert scale from 1 = *Not like me at all* to 7 = *Very much like me*. Higher scores correspond to higher universalism.

Connectedness to nature was assessed with the IINS, a graphical instrument developed by Kleespies et al. (2021). Two circles, respectively indicating the self and the natural world, are shown as gradually interconnected, representing the degree of perceived connection with nature. Respondents are thus asked to point out the extent to which they feel a sense of oneness with nature by choosing between seven response alternatives. Higher values suggest a stronger connection.

Pro-environmental behaviors were measured by a 6-item scale developed ad hoc. The single-factor scale requires respondents to self-assess the frequency of specific daily environmentally-friendly behaviors across various domains (e.g., "I turn off the lights when I leave a room"). The six items ($\omega = .86$) were designed on a Likert-type scale ranging from 1 (*never*) to 7 (*always*).

TPM Vol. 31, No. 3, September 2024
345-361
© 2024 Cises

Robba, M., Iannello, P.,
& Sorgente, A.
Italian adaptation of the GREEN scale

Data Analysis

On the first subsample (n = 503), descriptive statistics for the six items of the scale were analyzed to evaluate data normality and estimate means and standard deviations. Skewness and kurtosis were controlled, to check that the items fell within the suggested cutoffs ($\pm$ 1; Muthén & Kaplan, 1985). Subsequently, after evaluating the suitability of data by estimating the Bartlett's test of sphericity (p < .05; Bartlett, 1950) and the Kaiser-Meyer-Olkin test (KMO > .50; Kaiser, 1974), an EFA was performed in SPSS 29 using the principal axis factor extraction method. Besides observing the factorial structure, the quality of the items has been evaluated considering factor loadings (λ > .40; Howard, 2016), extracted communalities, and the corrected item-total correlations.

To collect validity evidence for the Italian version of the GREEN scale, different analyses were performed on the second subsample (n = 499). Indeed, following the contemporary view of validity theoretical framework (Hubley & Zumbo, 2011), construct validity of the GREEN scores was demonstrated by testing several types of validity evidence.

As for score structure evidence, a confirmatory factor analysis (CFA) was performed with the maximum likelihood robust (MLR) estimation method, to confirm the single-factor structure of the scale. For this purpose, different indices were considered to evaluate the goodness of fit of the model: the root-mean-square error of approximation (RMSEA), the comparative fit index (CFI), and the standardized root-mean-square residual (SRMR) were estimated. Confidence intervals (90%) of RMSEA and χ^2 significance test were also checked, although the latter is largely influenced by sample size (Bentler & Bonett, 1980). Hence, the relative fit index (χ^2/df) was controlled as well. RMSEA and SRMR scores lower than .08 indicate a good fit, while a CFI index higher than .90 is considered acceptable (Marsh et al., 2004). Finally, χ^2/df values < 2 suggest an adequate fit (Marsh & Hocevar, 1985). Once the mono-dimensional structure of the scale was confirmed, we evaluated the internal consistency of the GREEN scores (reliability evidence) estimating the omega coefficient (ω; McDonald, 1999) and the average variance extracted (AVE; Peterson et al., 2020). Commonly recommended acceptance thresholds are .70 for omega and .50 for AVE (Fornell & Larcker, 1981). Subsequently, the generalizability of the scale was tested by assessing measurement invariance across genders, to verify whether the factor structure was equivalent between males and females. MLR was adopted as the estimation method. Multigroup measurement invariance analyses consist of four steps (Vandenberg & Lance, 2000). Configural invariance evaluates whether the factor structure of the scale, the number of factors, and the item-factor correspondence are the same across groups. Secondly, through weak invariance, the similarity of factor loadings between groups was evaluated. Strong invariance tests the equivalence of item intercepts, while strict invariance assesses the equivalence of residual errors (Brown et al., 2017). The four models of invariance were compared adopting the free baseline approach (Stark et al., 2006). Specifically, the weak model was compared with the configural model, and the strong model was compared with the weak model. Finally, strict and strong models were compared. To check measurement invariance, CFI and RMSEA differences between models were evaluated, because χ^2 significance test is biased by sample size. As suggested by Chen (2007), a decrement of the CFI index higher than .01 suggests a meaningful decline of the model fit, as well as a decrease higher than .015 for the RMSEA index. After verifying that the scale was comparable between male and female participants (i.e., sufficient levels of measurement invariance), we compared the latent factor mean levels of GREEN across gender groups to test whether the differences expected according to the literature were detectable using the GREEN scale as well (known groups evidence). We expected females to have a higher level of biospheric values as suggested by Sargisson et al. (2020).

To check for convergent and divergent evidence, the scale was evaluated for association with other variables. Specifically, to collect convergent evidence a correlation model with latent indicators was performed,

considering both the GREEN scale and UNN subscale from the PVQ-RR (Schwartz & Cieciuch, 2022). Instead, discriminant evidence was tested by performing a second correlation model with the GREEN scale and the IINS graphical measure. Latent indicators were considered in the model as well. Finally, as for criterion-related evidence, the predictive validity of the GREEN scale was evaluated by performing a regression model with latent indicators. The capability of the scale scores to predict the scores of the ad hoc measure for pro-environmental behaviors was thus tested. The goodness of fit for each of the three SEM was evaluated considering the following indices: χ^2 significance test, RMSEA, CFI, and SRMR.

RESULTS

First Subsample: Item Analysis, EFA, Reliability

The first subsample was composed of 503 participants. Both skewness and kurtosis reported a normal distribution of the six items (with skewness values ranging from −.82 to −.60 and kurtosis values ranging from −.09 to .50). After checking for normality distribution, an EFA was performed. The Barlett's test of sphericity and the KMO measure of sampling adequacy provided satisfactory results. The Bartlett's test was significant: $\chi^2 = 1746.76$ ($df = 15$, $p < .001$) and the KMO was .90. The eigenvalue criteria suggested a single-factor structure for the scale. Factor loadings exceeded the suggested cut-off, with values ranging from .74 to .81. As reported in Table 2, extracted communalities were higher than .50 and the corrected item-total correlation values were all above .70. Finally, the single-factor structure explained the 61.62% of the total variance. The composite reliability of the scale was also good ($\omega = .91$).

TABLE 2
Descriptive statistics and items factor loading for the EFA

	M	SD	Skewness	Kurtosis	Factor loading	Extracted communality	Item-total correlation
Item 1	5.29	1.44	−.78	.50	.78	.61	.73
Item 2	5.09	1.58	−.77	.23	.81	.66	.77
Item 3	4.81	1.62	−.60	−.09	.79	.63	.75
Item 4	5.21	1.57	−.82	.38	.74	.55	.70
Item 5	5.24	1.41	−.77	.49	.78	.60	.73
Item 6	4.83	1.56	−.64	.11	.80	.64	.76
Explained variance (%)	61.62						

Note. EFA = exploratory factor analysis; M = mean; SD = standard deviation. Corrected item-total correlations are reported.

Second Subsample: Collecting Validity Evidence

Score Structure and Reliability Evidence

The CFA performed on a subsample of 499 respondents indicated that the single-factor theoretical model had good fit indices: $\chi^2 = 13.50$, $df = 9$, $p = .141$; $\chi^2/df = 1.50$; RMSEA = .032, 90% CI [.000, .064];

TPM Vol. 31, No. 3, September 2024
345-361
© 2024 Cises

Robba, M., Iannello, P.,
& Sorgente, A.
Italian adaptation of the GREEN scale

CFI = .992; SRMR = .021. The goodness of the model was also confirmed by the high factor loadings of the items (λ values ranging from .72 to .83), which resulted significant for $p < .001$ (see Table 3). Finally, internal consistency scores ($\omega = .90$; AVE = .60) suggested that the scale was reliable as well.

TABLE 3
Descriptive statistics and items factor loading for the CFA

	M	SD	Skewness	Kurtosis	Factor loading
Item 1	5.22	1.39	−.65	.41	.82***
Item 2	5.05	1.45	−.59	.11	.83***
Item 3	4.77	1.55	−.47	−.23	.79***
Item 4	5.22	1.51	−.82	.42	.74***
Item 5	5.25	1.44	−.80	.57	.72***
Item 6	5.00	1.46	−.63	.14	.77***

Note. CFA = confirmatory factor analysis; M = mean; SD = standard deviation.
***$p < .001$.

Generalizability and Known Groups Evidence

The results of the multigroup CFAs showed a good fit of the model along the four steps of the measurement invariance (i.e., configural, weak, strong, and strict invariance). Specifically, the difference in the χ^2 statistic was always nonsignificant, and differences in CFI and RMSEA indices when comparing models were below the suggested cutoffs (i.e., ΔCFI < .01; ΔRMSEA < .015), thus confirming that the factor structure of the scale was equivalent across genders. Table 4 indicates models fit along the four measurement invariance models.

TABLE 4
Fit indices for the invariant model

	χ^2	df	p	χ^2/df	RMSEA [90% CI]	CFI	SRMR	$\Delta\chi^2$	Δdf	p	ΔCFI	ΔRMSEA
Configural	25.04	18	.124	1.39	.040 [.000, .074]	.988	.031					
Weak	30.85	23	.127	1.34	.037 [.000, .068]	.987	.046	4.59	5	.467	−.001	−.003
Strong	36.09	28	.140	1.29	.034 [.000, .063]	.986	.045	3.74	5	.587	−.001	−.003
Strict	47.19	34	.066	1.39	.039 [.000, .065]	.978	.071	10.07	6	.122	−.008	.005

Note. χ^2 = chi-square; df = degree of freedom; RMSEA = root-mean-square error of approximation; CI = confidence interval; CFI = comparative fit index; SRMR = standardized root-mean-square residual.

The comparison to detect differences on the scores of the scale across genders turned out to be nonsignificant ($p = .113$). In particular, the mean level of biospheric values reported by women was higher than the one reported by men when looking at both latent scores ($M_{\text{female}} = .00$; $M_{\text{male}} = −.15$) and observed scores ($M_{\text{female}} = 5.18$, $SD = 1.23$; $M_{\text{male}} = 4.99$, $SD = 1.17$), but this difference was not statistically significant.

TPM Vol. 31, No. 3, September 2024
345-361
© 2024 Cises

Robba, M., Iannello, P.,
& Sorgente, A.
Italian adaptation of the GREEN scale

Convergent, Divergent, and Criterion-Related Evidence

Convergent, divergent, and criterion-related evidence was tested by performing three different SEM using latent variables as indicators to remove the measurement errors from these validity estimates.

Convergent evidence. The correlation model between biospheric values and universalism reported good fit indices: $\chi^2 = 41.79$, $df = 26$, $p = .026$; $\chi^2/df = 1.61$; RMSEA = .035, 90% CI [.012, .054]; CFI = .986; SRMR = .026. Furthermore, biospheric values resulted significantly and positively associated with universalism ($r = .86$, $p < .001$). Pearson's correlation was in the expected direction and with a strong effect size, confirming convergent evidence.

Discriminant evidence. Fit indices indicated that the correlation model between biospheric values and connectedness to nature was good: $\chi^2 = 19.09$, $df = 14$, $p = .162$; $\chi^2/df = 1.36$; RMSEA = .027, 90% CI [.000, .054]; CFI = .993; SRMR = .021. The estimated model showed a significant and positive relationship between the two variables ($r = .45$, $p < .001$). As expected, the correlation coefficient between biospheric values and connectedness to nature (i.e., discriminant evidence) was weaker than that between biospheric values and universalism (i.e., convergent evidence). This finding confirmed the discriminant evidence of the scale.

Criterion-related evidence. The full regression model with biospheric values predicting pro-environmental behaviors resulted adequate: $\chi^2 = 92.91$, $df = 53$, $p = .001$; $\chi^2/df = 1.75$; RMSEA = .039, 90% CI [.025, .052]; CFI = .975; SRMR = .046. Values were found to be a strong and positive predictor of pro-environmental behaviors ($\beta = .52$, $p < .001$), thus highlighting the presence of criterion-related evidence. Furthermore, biospheric values explained almost a quarter of the total variance of pro-environmental behaviors ($R^2 = .26$).

DISCUSSION

The present paper aimed to validate the Italian version of the GREEN scale, developed by Haws and colleagues (2014) following the guidelines of the contemporary view of validity (Hubley & Zumbo, 2011). This approach has two major advantages. First, it does not rely on a single type of validity to validate an instrument, but, instead, collects multiple pieces of evidence to prove that the instrument under study does indeed measure the construct it is expected to measure. Second, this evidence is gathered using state-of-the-art statistical techniques that allow measurement error to be removed from the estimates. In summary, conclusions drawn using the contemporary view of validity approach are more reliable.

Findings supported the unidimensionality of the scale and its factorial invariance across genders. Indeed, the EFA replicated the original factorial structure of the scale and the CFA provided good fit indices for the theoretical model. The measurement invariance across genders also supported that the single-factor measure of biospheric values was equivalent between males and females. However, we did not find the expected gender differences in the average scores of biospheric values. As stated in the Introduction, the literature on this topic is still limited and not univocal. Indeed, while gender differences in pro-environmental behaviors are a well-known phenomenon, the few previous studies on biospheric values reported only a marginal variability between men and women (e.g., Sargisson et al., 2020). This could suggest that the gender difference is present at a behavioral level, but less marked in its antecedents, such as values. We also argue that, nowadays, climate change and environmental issues are gaining more and more emphasis in our lives. Hence, the marginal gender differences found in the present study might also be explained by the increase of public attention on the phenomenon.

TPM Vol. 31, No. 3, September 2024
345-361
© 2024 Cises

Robba, M., Iannello, P.,
& Sorgente, A.
Italian adaptation of the GREEN scale

Furthermore, convergent, discriminant, and criterion validity evidence were supported, because biospheric values were significantly related to universalism (UNN from PVQ-RR; Schwartz & Cieciuch, 2022), connectedness to nature (IINS; Kleespies et al., 2021), and pro-environmental behaviors, respectively. Specifically, the high correlation between the scores of the GREEN scale (i.e., biospheric values) and UNN subscale (i.e., universalism) suggested that the two instruments measure the same psychological construct. On the other hand, the relationship between biospheric values and connectedness to nature was significant, though with a smaller effect size compared to the correlation between biospheric values and universalism. This highlighted that the GREEN scale and the IINS measure two different and separate constructs. Finally, biospheric values significantly predicted daily environmentally-friendly behaviors, confirming the idea that people's beliefs and principles may explain people's behaviors (e.g., Wang et al., 2021).

In brief, the GREEN scale proved to be a valid and reliable measure for biospheric values in the Italian context. In agreement with the contemporary view of validity, we can conclude that when this scale is used in a context (e.g., Italian) and with a sample (e.g., people over 18 years old) similar to the one adopted in the current study, users can rely on test scores to make inferences about biospheric values. When, instead, the measure is adopted in a different context or with a different sample, test users should first collect new evidence to be sure that the instrument measures what it is supposed to measure.

LIMITATIONS

This study does not lack limitations. First, considering its cross-sectional nature, any inference on the causal relationship between biospheric values and environmentally-friendly behaviors should be avoided. Future research should adopt a longitudinal design to overcome this limit and test the predictive version other than the concurrent version of the criterion-related validity. Moreover, self-reported measures for pro-environmental behaviors should be somehow biased by social desirability and other response biases. Behavioral data should also be considered when investigating the role of biospheric values in shaping individuals' actions toward the natural environment. A third methodological limit refers to the absence of a multitrait-multimethod matrix (MTMM) analysis usually adopted to assess the convergent and discriminant evidence of the scale (Campbell & Fiske, 1959). Because in our study all the variables were measured using the same method (i.e., self-report assessment) we could not collect this more complete evidence of convergent/discriminant validity, but we had to test them separately using a single correlation coefficient. Finally, the known groups evidence of the scale was not supported (i.e., the expected difference between women and men was present at a descriptive level but not statistically significant). Future studies should better investigate gender differences in biospheric values, considering that the current literature is still limited and discordant.

THEORETICAL AND PRACTICAL IMPLICATIONS

The Italian version of the GREEN scale showed good psychometric properties, proving to be a valid and reliable instrument for measuring concern in environmental issues. Furthermore, considering that it was recently validated and its brevity it can be easily integrated into questionnaires and used in a variety of settings and contexts. This study contributes to the literature of both biospheric values and pro-environmental behaviors, by showing a strong relationship between the two variables. A second theoretical implication concerned the relationship between biospheric values and connectedness to nature. Specifically, the

TPM Vol. 31, No. 3, September 2024
345-361
© 2024 Cises

Robba, M., Iannello, P.,
& Sorgente, A.
Italian adaptation of the GREEN scale

theoretical distinction between the two constructs was empirically supported. Sustainability and climate change issues are becoming increasingly salient nowadays and an Italian adaptation of this instrument can benefit future studies investigating individuals' concern toward environmental conservation. Having a psychometric scale measuring biospheric values in the Italian context allows researchers to better understand how values could shape environmentally-friendly behaviors and sustainable consumption preferences, or investigate which psychological barriers could increase the so-called value-behavior gap. Such research topics could have practical implications in helping policymakers develop environmental education programs or behavioral interventions aimed to promote pro-environmental behaviors among citizens. This instrument could also be applied in market research, for instance to understand different types of consumers and subsequently develop ad hoc advertising campaigns.

REFERENCES

Bailey, A. A., Mishra, A. S., & Tiamiyu, M. F. (2018). Application of GREEN scale to understanding US consumer response to green marketing communications. *Psychology & Marketing*, *35*(11), 863-875. https://doi.org/10.1002/mar.21140

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, *3*(2), 77-5. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x

Bartoszczuk, P., Singh, U., & Rutkowska, M. (2022). An empirical analysis of renewable energy contributions considering green consumer values — A case study of Poland. *Energies*, *15*(3), Article 1027. https://doi.org/10.3390/en15031027

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588-606. https://doi.org/10.1037/0033-2909.88.3.588

Bonera, M., Codini, A. P., & Miniero, G. (2020). The great Millennials' trouble: Leading or confused green generation? An Italian insight. *Italian Journal of Marketing*, *2020*(4), 289-308. https://doi.org/10.1007/s43039-020-00015-4

Briscoe, M. D., Givens, J. E., Hazboun, S. O., & Krannich, R. S. (2019). At home, in public, and in between: gender differences in public, private and transportation pro-environmental behaviors in the US Intermountain West. *Environmental Sociology*, *5*(4), 374-392. https://doi.org/10.1080/23251042.2019.1628333

Brown, G. T. L., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: Identifying and understanding non-invariance. *International Journal of Research & Method in Education*, *40*(1), 66-90. https://doi.org/10.1080/1743727x.2015.1070823

Bouman, T., Steg, L., & Kiers, H. A. L. (2018). Measuring values in environmental research: A test of an Environmental Portrait Value Questionnaire. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00564

Bouman, T., Verschoor, M., Albers, C. J., Böhm, G., Fisher, S., Poortinga, W., Whitmarsh, L., & Steg, L. (2020). When worry about climate change leads to climate action: How values, worry and personal responsibility relate to various climate actions. *Global Environmental Change*, *62*, Article 102061. https://doi.org/10.1016/j.gloenvcha.2020.102061

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81-105. https://doi.org/10.1037/h0046016

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464-504. https://doi.org/10.1080/10705510701301834

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302. https://doi.org/10.1037/h0040957

De Groot, J. I. M., & Steg, L. (2007). Value orientations and environmental beliefs in five countries. *Journal of Cross-cultural Psychology*, *38*(3), 318-332. https://doi.org/10.1177/0022022107300278

De Groot, J. I. M., & Steg, L. (2009). Mean or green: Which values can promote stable pro-environmental behavior? *Conservation Letters*, *2*(2), 61-66. https://doi.org/10.1111/j.1755-263x.2009.00048.x

De Groot, J. I. M., & Steg, L. (2010). Relationships between value orientations, self-determined motivational types and pro-environmental behavioural intentions. *Journal of Environmental Psychology*, *30*(4), 368-378. https://doi.org/10.1016/j.jenvp.2010.04.002

De Groot, J. I. M., & Thøgersen, J. (2018). Values and pro-environmental behaviour. In L. Steg & J. I. M. De Groot (Eds.), *Environmental psychology: An introduction* (pp. 167-178). Wiley. https://doi.org/10.1002/9781119241072.ch17

Dikici, Z. Y., Çakrak, M., & Demirci, E. (2022). Green consumption values, social appreciation, and purchasing behavior. *Sustainability and Climate Change*, *15*(3), 189-199. https://doi.org/10.1089/scc.2022.0003

Dupont, J., Harms, T., & Fiebelkorn, F. (2022). Acceptance of cultured meat in Germany — Application of an extended theory of planned behaviour. *Foods*, *11*(3), Article 424. https://doi.org/10.3390/foods11030424

Essiz, O., & Mandrik, C. (2021). Intergenerational influence on sustainable consumer attitudes and behaviors: Roles of family communication and peer influence in environmental consumer socialization. *Psychology & Marketing*, *39*(1), 5-26. https://doi.org/10.1002/mar.21540

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *JMR, Journal of Marketing Research*, *18*(1), 39-50. https://doi.org/10.1177/002224378101800104

Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, *6*(4), 304-312. https://doi.org/10.1037/1040-3590.6.4.304

Haws, K. L., Winterich, K. P., & Naylor, R. W. (2014). Seeing the world through green-tinted glasses: GREEN consumption values and responses to environmentally friendly products. *Journal of Consumer Psychology*, *24*(3), 336-354. https://doi.org/10.1016/j.jcps.2013.11.002

Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human-Computer Interaction*, *32*(1), 51-62. https://doi.org/10.1080/10447318.2015.1087664

Hubley, A. M., & Zumbo, B. D. (2011). Validity and the consequences of test interpretation and use. *Social Indicators Research*, *103*(2), 219-230. https://doi.org/10.1007/s11205-011-9843-4

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36. https://doi.org/10.1007/BF02291575

Katz-Gerro, T., Greenspan, I., Handy, F., & Lee, H.-Y. (2017). The relationship between value types and environmental behaviour in four countries: Universalism, benevolence, conformity and biospheric values revisited. *Environmental Values*, *26*(2), 223-249. https://www.jstor.org/stable/44202483

Kleespies, M. W., Braun, T., Dierkes, P. W., & Wenzel, V. (2021). Measuring connection to nature — A illustrated extension of the inclusion of nature in self scale. *Sustainability*, *13*(4), Article 1761. https://doi.org/10.3390/su13041761

Klein, F. F., Emberger-Klein, A., & Menrad, K. (2020). Indicators of consumers' preferences for bio-based apparel: A German case study with a functional rain jacket made of bioplastic. *Sustainability*, *12*(2), Article 675. https://doi.org/10.3390/su12020675

Lengieza, M. L., & Swim, J. K. (2021). The paths to connectedness: A review of the antecedents of connectedness to nature. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.763231

Macht, J., Klink-Lehmann, J., & Venghaus, S. (2023). Eco-friendly alternatives to food packed in plastics: German consumers' purchase intentions for different bio-based packaging strategies. *Food Quality and Preference*, *109*, Article 104884. https://doi.org/10.1016/j.foodqual.2023.104884

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320-341. https://doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin, 97*(3), 562-582. https://doi.org/10.1037/0033-2909.97.3.562

Martin, C., & Czellar, S. (2017). Where do biospheric values come from? A connectedness to nature perspective. *Journal of Environmental Psychology*, *52*, 56-68. https://doi.org/10.1016/j.jenvp.2017.04.009

Mazzocchi, C., Orsi, L., & Sali, G. (2021). Consumers' attitudes for sustainable mountain cheese. *Sustainability*, *13*(4), Article 1743. https://doi.org/10.3390/su13041743

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, *18*(2), 5-11. https://doi.org/10.3102/0013189x018002005

TPM Vol. 31, No. 3, September 2024
345-361
© 2024 Cises

Robba, M., Iannello, P.,
& Sorgente, A.
Italian adaptation of the GREEN scale

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *The American Psychologist*, *50*(9), 741-749. https://doi.org/10.1037/0003-066x.50.9.741

Milfont, T. L., & Sibley, C. G. (2016). Empathic and social dominance orientations help explain gender differences in environmentalism: A one-year Bayesian mediation analysis. *Personality and Individual Differences*, *90*, 85-88. https://doi.org/10.1016/j.paid.2015.10.044

Muthén, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *The British Journal of Mathematical and Statistical Psychology*, *38*(2), 171-189. https://doi.org/10.1111/j.2044-8317.1985.tb00832.x

Paço, A., Shiel, C., & Alves, H. (2019). A new model for testing green consumer behaviour. *Journal of Cleaner Production*, *207*, 998-1006. https://doi.org/10.1016/j.jclepro.2018.10.105

Pegan, G., Del Missier, F., & De Luca, P. (2023). Antecedents of green purchase choices: Towards a value-oriented model. *Journal of Cleaner Production*, *399*, Article 136633. https://doi.org/10.1016/j.jclepro.2023.136633

Perlaviciute, G., & Steg, L. (2015). The influence of values on evaluations of energy alternatives. *Renewable Energy*, *77*, 259-267. https://doi.org/10.1016/j.renene.2014.12.020

Peterson, R. A., Kim, Y., & Choi, B. (2020). A meta-analysis of construct reliability indices and measurement model fit metrics. *Methodology: European Journal of Research Methods for the Behavioral & Social Sciences*, *16*(3), 208-223. https://doi.org/10.5964/meth.2797

Rahman, I., & Reynolds, D. (2016). Predicting green hotel behavioral intentions using a theory of environmental commitment and sacrifice for the environment. *International Journal of Hospitality Management*, *52*, 107-116. https://doi.org/10.1016/j.ijhm.2015.09.007

Ribeiro, M. A., Seyfi, S., Elhoushy, S., Woosnam, K. M., & Patwardhan, V. (2023). Determinants of generation Z pro-environmental travel behaviour: The moderating role of green consumption values. *Journal of Sustainable Tourism*, 1-21. https://doi.org/10.1080/09669582.2023.2230389

Risitano, M., Romano, R., La Ragione, G., & Quintano, M. (2023). Analysing the impact of green consumption values on brand responses and behavioural intention. *Business Ethics, the Environment & Responsibility*, *32*(3), 1096-1112. https://doi.org/10.1111/beer.12543

Ruepert, A. M., Keizer, K., & Steg, L. (2017). The relationship between corporate environmental responsibility, employees' biospheric values and pro-environmental behaviour at work. *Journal of Environmental Psychology*, *54*, 65-78. https://doi.org/10.1016/j.jenvp.2017.10.006

Sargisson, R. J., De Groot, J. I. M., & Steg, L. (2020). The relationship between sociodemographics and environmental values across seven European countries. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.02253

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*(4), 457-475. https://doi.org/10.1177/0022022105275962

Schultz, P. W., & Zelezny, L. (1999). Values as predictors of environmental attitudes: Evidence for consistency across 14 countries. *Journal of Environmental Psychology*, *19*(3), 255-265. https://doi.org/10.1006/jevp.1999.0129

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, *25*, 1-65. https://doi.org/10.1016/S0065-2601(08)60281-6

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *The Journal of Social Issues*, *50*(4), 19-45. https://doi.org/10.1111/j.1540-4560.1994.tb01196.x

Schwartz, S. H., & Cieciuch, J. (2022). Measuring the refined theory of individual values in 49 cultural groups: Psychometrics of the revised Portrait Value Questionnaire. *Assessment*, *29*(5), 1005-1019. https://doi.org/10.1177/1073191121998760

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.-E., Demirutku, K., Dirilen-Gumus, O., & Konty, M. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, *103*(4), 663-688. https://doi.org/10.1037/a0029393

Sorgente, A., & Zumbo, B. D. (in press). The alphas and omegas of validity and reliability: Contemporary advances in evaluating and selecting instruments for quantitative research with emerging adults. In A. Sorgente, R. Vosylis, S. Claxton, & J. Schwab (Eds.), *Flourishing as a scholar: Research methods for the study of emerging adulthood*. Oxford University Press.

Stahl, F. F., Emberger-Klein, A., & Menrad, K. (2021). Consumer preferences in Germany for bio-based apparel with low and moderate prices, and the influence of specific factors in distinguishing between these groups. *Frontiers in Sustainability*, *2*. https://doi.org/10.3389/frsus.2021.624913

Stark, S., Chernyshenko, O. S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *The Journal of Applied Psychology*, *91*(6), 1292-1306. https://doi.org/10.1037/0021-9010.91.6.1292

Steg, L., Perlaviciute, G., van der Werff, E., & Lurvink, J. (2014). The significance of hedonic values for environmentally relevant attitudes, preferences, and actions. *Environment and Behavior*, *46*(2), 163-192. https://doi.org/10.1177/0013916512454730

van der Werff, E., Steg, L., & Keizer, K. (2013). The value of environmental self-identity: The relationship between biospheric values, environmental self-identity and environmental preferences, intentions and behaviour. *Journal of Environmental Psychology*, *34*, 55-63. https://doi.org/10.1016/j.jenvp.2012.12.006

van Doorn, J., & Verhoef, P. C. (2015). Drivers of and barriers to organic purchase behavior. *Journal of Retailing*, *91*(3), 436-450. https://doi.org/10.1016/j.jretai.2015.02.003

van Widenfelt, B. M., Treffers, P. D. A., de Beurs, E., Siebelink, B. M., & Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child and Family Psychology Review*, *8*(2), 135-147. https://doi.org/10.1007/s10567-005-4752-1

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4-70. https://doi.org/10.1177/109442810031002

Vicente-Molina, M. A., Fernández-Sáinz, A., & Izagirre-Olaizola, J. (2013). Environmental knowledge and other variables affecting pro-environmental behaviour: Comparison of university students from emerging and advanced countries. *Journal of Cleaner Production*, *61*, 130-138. https://doi.org/10.1016/j.jclepro.2013.05.015

Vicente-Molina, M. A., Fernández-Sainz, A., & Izagirre-Olaizola, J. (2018). Does gender make a difference in pro-environmental behavior? The case of the Basque Country University students. *Journal of Cleaner Production*, *176*, 89-98. https://doi.org/10.1016/j.jclepro.2017.12.079

Wang, X., van der Werff, E., Bouman, T., Harder, M. K., & Steg, L. (2021). I am vs. We are: How biospheric values and environmental identity of individuals and groups can influence pro-environmental behaviour. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.618956

Wiliam, D. (2014). What do teachers need to know about the new standards for educational and psychological testing? *Educational Measurement Issues and Practice*, *33*(4), 29-30. https://doi.org/10.1111/emip.12051

Xia, W., & Li, L. M. W. (2023). Societal gender role beliefs moderate the pattern of gender differences in public- and private-sphere pro-environmental behaviors. *Journal of Environmental Psychology*, *92*, Article 102158. https://doi.org/10.1016/j.jenvp.2023.102158

Zumbo, B. D. (2005). Structural equation modeling and test validation. In B. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (pp. 1951-1958). John Wiley & Sons Ltd. https://doi.org/10.1002/0470013192.bsa654

Zumbo, B. D. (2006). Validity: Foundational issues and statistical methodology. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics, Vol. 26: Psychometrics* (pp. 45-79). Elsevier.

TPM®

APPENDIX

Original and translated items of the GREEN scale

Original items	Translated items
1. It is important to me that the products I use do not harm the environment	1. È importante per me usare prodotti che non causino danni all'ambiente
2. I consider the potential environmental impact of my actions when making many of my decisions	2. Quando prendo decisioni pongo attenzione all'impatto che queste possono avere sull'ambiente
3. My purchase habits are affected by my concern for our environment	3. Le mie abitudini d'acquisto sono influenzate dalla mia preoccupazione per l'ambiente
4. I am concerned about wasting the resources of our planet	4. Sono preoccupato di sprecare le risorse a disposizione del pianeta
5. I would describe myself as environmentally responsible	5. Mi descriverei come una persona responsabile verso l'ambiente
6. I am willing to be inconvenienced in order to take actions that are more environmentally friendly	6. Sono disposto a rinunciare alla comodità per adottare comportamenti più rispettosi verso l'ambiente