Consumer perception of brand equity based on environmental sustainability – development and validation of a scale

Percepção do consumidor de brand equity com base na sustentabilidade ambiental - desenvolvimento e validação de uma escala

DOI: 10.55905/oelv21n6-087

Recebimento dos originais: 26/05/2023
Aceitação para publicação: 26/06/2023

Luiz Diego Vidal Santos
Doctoral Student in Intellectual Property Sciences
Institution: Universidade Federal de Sergipe (UFS)
Address: Av. Marechal Rondon, S/N, Jardim Rosa Elze, São Cristóvão - SE, CEP: 49100-000
E-mail: vidal.center@academico.ufs.br

Francisco Sandro Rodrigues Holanda
PhD in Agronomy
Institution: Universidade Federal de Sergipe (UFS)
Address: Av. Marechal Rondon, S/N, Jardim Rosa Elze, São Cristóvão - SE, CEP: 49100-000
E-mail: fholanda@academico.ufs.br

Catuxe Varjão de Santana Oliveira
Master in Computer Science
Institution: Instituto Federal de Sergipe (IFS)
Address: Av. Eng. Gentil Tavares, 1166, Getúlio Vargas, Aracaju - SE, CEP: 49055-260
E-mail: catuxe@gmail.com

Pedro Vinícius Bertulino de Menezes
Agronomic Engineering Undergraduate
Institution: Universidade Federal de Sergipe (UFS)
Address: Av. Marechal Rondon, S/N, Jardim Rosa Elze, São Cristóvão - SE, CEP: 49100-000
E-mail: pedroviniciusbm20@gmail.com
Ana Paula Schervinski Villwock
PhD in Rural Extension
Institution: Universidade Federal de Sergipe (UFS)
Address: Av. Marechal Rondon, S/N, Jardim Rosa Elze, São Cristóvão - SE, CEP: 49100-000
E-mail: ana.agronomia@gmail.com

Alceu Pedrotti
PhD in Soil Science
Institution: Universidade Federal de Sergipe (UFS)
Address: Av. Marechal Rondon, S/N, Jardim Rosa Elze, São Cristóvão - SE, CEP: 49100-000
E-mail: alceupedrotti@gmail.com

ABSTRACT
Brands gradually became the core dimension and strategic asset of branding for organizations of all sizes, and today many companies adopt various forms of green marketing activities as part of their strategies. The objective of this study was to develop and validate the psychometric properties of precision and validity of a scale for assessing brand equity based on environmental sustainability. The scale was validated by adopting the standards described in APA Standard. Eight companies were selected from those developing their sustainable business as samples for validation tests. Data were applied to 263 potential consumers. A scale with 2 items was built and validated, divided into four dimensions: Brand quality perception BQP (6 questions); Strategic brand positioning SBP (10 questions); Willingness to buy WTB (3 questions) and Retail Innovation RI (3 questions). The instrument items showed good internal consistency regarding their domains. Matrix is interpreted in 4 dimensions. Factorial analysis assumed few cross-loadings, corroborating the initial assumptions of the adapted instruments. The instrument obtained good internal fit indices and external measures. Finally, Item response theory (IRT) analysis showed no floor or ceiling effect; however, for the DIF data, the scale invariably works for older and younger people for almost all items except RI 04 item. This article fills the gap in literature by developing a scale to assess consumer perception of brand equity and environmental sustainability, improving the understanding of the customer experience with brand equity, exploring various dimensions of consumer perception of sustainable brands.

Keywords: forest sustainability, sustainable, scale validation, psychometric instrument, amazon rainforest.

RESUMO
As marcas tornaram-se gradualmente a dimensão central e o ativo estratégico do branding para organizações de todos os tamanhos, e hoje muitas empresas adotam várias formas de atividades de marketing verde como parte de suas estratégias. O objetivo deste estudo foi desenvolver e validar as propriedades psicométricas de precisão e validade de uma escala de avaliação de brand equity com base na sustentabilidade ambiental. A escala foi
validada adotando os padrões descritos na Norma APA. Oito empresas foram selecionadas entre aquelas que desenvolvem seus negócios sustentáveis como amostras para testes de validação. Os dados foram aplicados a 263 potenciais consumidores. Foi construída e validada uma escala com 2 itens, divididos em quatro dimensões: Percepção da qualidade da marca BQP (6 questões); Posicionamento estratégico da marca SBP (10 questões); Vontade de comprar WTB (3 questões) e Retail Innovation RI (3 questões). Os itens do instrumento apresentaram boa consistência interna quanto aos seus domínios. A matriz é interpretada em 4 dimensões. A análise fatorial assumiu poucas cargas cruzadas, corroborando as suposições iniciais dos instrumentos adaptados. O instrumento obteve bons índices de ajuste interno e medidas externas. Finalmente, a análise da teoria de resposta ao item (TRI) não mostrou efeito de piso ou teto; no entanto, para os dados DIF, a escala invariavelmente funciona para pessoas mais velhas e mais jovens para quase todos os itens, exceto o item RI 04. Este artigo preenche a lacuna na literatura ao desenvolver uma escala para avaliar a percepção do consumidor sobre brand equity e sustentabilidade ambiental, melhorando a compreensão da experiência do cliente com brand equity, explorando várias dimensões da percepção do consumidor sobre marcas sustentáveis.

Palavras-chave: sustentabilidade florestal, sustentável, validação de escala, instrumento psicométrico, floresta amazônica.

1 INTRODUCTION

Brand equity and sustainability have been constantly discussed in the academic scenario. (Moritz et al., 2018) states that the two fundamentals should be seen as symbiotic when planning to change the image of the business with a focus on innovation in the current setlist. Other studies reinforce that sustainability is an antecedent and basic factor for the brand's heavy equity (Arrigo, 2018; Gonzalez-Lafaysse & Lapassouse-Madrid, 2016; Ray & Sharma, 2020). Recent research points to a link between these variables that can contribute to an organization's economic and sustainable advantages (Marín-García et al., 2020).

As the main asset for the positive perception of the consumer, brands can be understood as labels, symbols and descriptions that involve meanings and associations, resulting in a flexible contract with the customer in the form of providing a product or service, different signs of products, companies or services (Wood, 2007). As a result, they have become an important economic asset for their bearer, determining how much the consumer accepts the brand's signatory company's product, significantly impacting its
commercial performance (Singh et al., 2020). In view of this, several companies emerge or, when consolidated, adopt sustainable policies in all spheres of contact with the consumer.

However, changes in the strategic characteristics of the new corporate image and the implementation of environmental innovations require a solid and strong engagement on the part of the company linked to the brand in its products and manufacturing processes and/or services of the companies (W. L. Lin et al., 2021). Since, currently with the ease of access to information, consumers have increasingly sought information about the sustainable standards adopted by the brands they consume (S. Y. Chen, 2019).

One of the ways to maintain competitive capacity and produce sustainably, maintaining the value of your products and services or even adding additional premium value, is the adoption of branding management based on consumer perception (Crader & Santomier, 2011). This management strategy seeks to associate the differentiation in sustainable production obtained through sustainable technologies to the points of contact perceived by consumers (Choi & Yi, 2018; Long & Lin, 2018; Marques et al., 2020). Obtaining data related to the consumer's perception of the brand's image can be conceived through the evaluation of first-order formative constructs (Aprile and Punzo, 2022; Cowan and Dai, 2014; Defrancesco et al. 2017). These formative constructs can be evaluated from the domain of consumer experiences, through the perception of brand image, product quality and strategic positioning of the brand. (J. L. Aaker, 1997; Figueiredo & Castro, 2019). All these constructs can be observed from image analyses.

Brand-related commercial images are defined as pre-existing understandings associated with new information contained in the brand's visual identity (Cowan & Dai, 2014). The concepts contained in the brand value directly related to the assets of the enterprise, once properly constructed, promote the consumer's affinity with the products by understanding positive and ethically acceptable sensory cues (Ulian et al., 2020). Previous research shows that images related to consumption increase attitudes and willingness to buy (Schlosser, 2003), forming concepts about products and their value and fair price (Chan et al., 2018).
The perception of quality is at a higher level of abstraction than any specific attribute for choosing and purchasing a brand's products. It can be defined as the customer's judgment of general excellence, estimating the brand's superiority over competitors (Konuk, 2018); it's a brand capability assessment—a general emotional assessment of a brand's performance relative to other brands (Y. S. Chen & Chang, 2013). The perception of quality is indissolubly intertwined with receiving functional/practical benefits that increase the value of the reward for purchasing the product, albeit with added premium value (Dash et al., 2021).

Studies show that the adoption of additional informational data from the manufacturing and environmental policies adopted by the company increases the perception of the amount embedded in the product, leading to a greater transfer of product attributes from the strong brand to the rebranded brand (Churchill Jr, 1979; Jestratijevic et al., 2022; Khandeparkar, 2014; Zhao, 2022). In this sense, the adoption of informational knowledge about environmental preservation policies adds value to the brand linked to the perceived quality of the product, converting it into a strong brand.

Brand positioning, on the other hand, can be understood through differentiation strategies, which aim to project the company's image, occupying a unique place in the mind of the target market and converting it into products that are trusted by consumers and brand propagators (Pratihari & Uzma, 2018). The notion of brand positioning has been seen as an offshoot of differentiation, through which researchers try to locate in the consumer's mind the relative position of a brand in relation to its competitors (McDowell, 2009; Wang & Horng, 2016).

However, many studies have shown that positioning is largely determined by the identity and image of a brand and its consequence in the formation of the consumer's imagination regarding the values it provides (Bulovic & Seric, 2021; le Roux & du Plessis, 2019; Wang, 2017).

The interplay between the brand concept and the formative values help to create strong, unique and positive associations that should resonate with the brand's positioning in a prominent place (Koch & Gyrd-Jones, 2019). Creating the points of difference is key to building consumer-friendly brand equity. It can increase the perception of brand
exclusivity (Ameyibor et al., 2022), a factor that significantly impact the perception of exclusivity of the business.

Purchase interest can be derived from the economic value a customer is willing to pay for their preferred brand versus a comparable/lesser brand for the same size/package quantity (Kimura et al., 2012; Konuk, 2019). It is one of the strongest indicators of brand loyalty and may be the most reasonable summation measure of total brand equity (Kataria & Saini, 2020; Netemeyer et al., 2004).

In this sense, we can define purchase interest as buyers' preferences for products manufactured and sold under conditions (Belanche et al., 2021) fair and substantially acceptable working conditions. Willingness to buy is one of the strongest indicators of brand loyalty and may be the most reasonable measure of overall brand equity (Balderjahn & Hüttel, 2019). Competitive advantages achieved through brand innovation or the relationship between innovation and market performance play an important role in the brand equity of sustainable companies (Medrano et al., 2020; Quaye & Mensah, 2019). The definition of branding innovation has been widely portrayed in different domains and academic perspectives. In a broader concept, it can be defined as “a new way of doing things, or a unique combination of production factors” (Schumpeter, 1980). However, there is a need for a new definition, due to changes in global business trends and the flow of innovative ideas partially from producer to consumer (Paswan et al., 2020).

The correct understanding of the latent variables that shape the nomological network to form the brand equity of environmentally sustainable brands in different cultures is relevant, as different populations have different social values and these can interfere with perception and choice (Zajkowski & Domańska, 2019).

In addition, observing the performance of different groups is fundamental for the equity of the instruments, as the assessments should not discriminate against individuals, or the sample assessed by factors that are irrelevant to the test construct. Research on the development of psychometric instruments suggests the creation of scales from a mix of original items and those adapted from other already validated research to provide an
assessment of the construct in all its possible dimensions (Boateng et al., 2018; Pandey & Chawla, 2014).

This approach provides a useful starting point for conforming to new contexts (Churchill Jr, 1979). In this sense, current research adopts the creation and adaptation of psychometric instruments through cross-cultural experimentation methods, considering the linguistic and cultural characteristics of the region to be studied based on other pre-existing studies. The cultural and linguistic adaptation of tools designed to assess consumer perception of sustainable products is an essential step to increase their usefulness and relevance.

Often, the development of a brand management model requires a polycentric understanding to reach enough consumer markets to generate the need to apply the tool internationally (H. Liu et al., 2022; Shrestha et al., 2020). To reliably collect and analyze instrument data, instruments in different target languages should cover the same concepts, although the text should be modified to reflect equivalent concepts in the target language (Hustvedt & Bernard, 2010).

The field of psychological perception analyzes, and cross-cultural adaptation has slowly reached consensus on how to effectively validate and use measures (Grassi et al., 2021). The fitting procedure brings together areas of expertise in different fields, including psychometrics and statistics (eg, structural equation modeling and item response models) (Clark & Watson, 2019; Y. Liu et al., 2019; McAllister et al., 2021; Mels et al., 2010). The cross-cultural adaptation process is essential when a scale is used in a different language, environment and time to reduce the risk of bias in a study (Herdman et al., 1998; Swami & Barron, 2019). In this study, the thematic adaptation of the items did not lead to discrepant adjustments of translation and cross-cultural adaptation, none of them harmful to the instrument.

The aim of this study was to investigate and develop the psychometric properties of accuracy and validity of a scale for assessing brand equity based on environmental sustainability in the Brazilian Amazon rainforest. The validation phases followed the psychometric instrument validation pattern. Standardized scale assessment tools should demonstrate similar interpretation of test items across cultures.
So, this research is organized as follows. First, we present the dimensions analyzed for the development of the scale. Next, we present the methodology adopted to validate and refine the proposed scale. Finally, we discuss the results obtained and the final considerations.

2 MATERIALS AND METHODS

2.1 OPERATIONALIZATION OF VARIABLES

Brand equity comprises a series of dimensions that the brand can develop, manage, and control. Different conceptual structures include different dimensions. Then, this research is considered a theoretical reference to the dimensions postulated by Aaker (1996): awareness, loyalty, perceived quality, and associations. The initial dimensions and findings from a scoping review were developed by cross-cultural translation (‘forward’) of selected items from the original instruments.

Two translators, one being a linguist and the other from the research area of the experiment, were selected since while translating a scale, several equivalences must be sought about the original, such as cultural, semantic, technical, content, and criteria (Guillemin et al., 1993). Furthermore, the current literature in the area highlights the need to avoid a literal translation of the items, a fact that can be avoided if they are only met through the variability of translators (Callegaro Borsa et al., 2012).

After the consensus of the translators, a structured scale initially contains 23 items to deal with brand equity of brands related to environmental sustainability in Amazon rainforest. The four dimensions proposed in this research were Brand quality perception — BQP (6 questions); Strategic Brand Positioning — SBP (10 questions); Willingness to buy — WTB (3 questions) and Retail inovation — RI (3 questions).

We investigated data from 7 brands related to sustainability that market their products/services, considering the equity of the sustainable brand. The brands are classified into six categories: cosmetics, tourism, clothing, accessories, and food. To characterize the perception of the brand, we gathered images from the website, images of the products, posts on social networks, and descriptive texts, as well as the prototyping of the e-commerce website, using data provided by them.
The total sample consisted of 263 potential consumers, aged between 16 and 64 (ages 19 to 30 = 40%; 31 to 64 = 60%). Participating interviewees were selected using the quota method to guarantee homogeneity regarding age and sex. Participants included consumers from all regions of Brazil. The survey was carried out from March to July 2022.

2.2 ANALYZED DIMENSIONS

The first dimension is composed of 10 (ten) transliterated questions from the developed scale by Grigorescu et al. (2019), which will analyze the interviewee's perception of the strategic positioning of the brand (SBP) in the Amazonian region of Brazil. Additionally, we considered the level of perception of the participants regarding the importance of the company/startup relationship with environmental sustainability in the region. Ten (10) with minor language modifications out of 21 (twenty-one) items of the original scale were used.

The second dimension presents five items and comprises the perception of product quality and the brand equity related to environmental sustainability initially proposed by each brand (BQP). This construct was evaluated using five pictures/items of the brand, five pictures of the store (digital), four pictures/items of the premium product (with a value above the average market), and five pictures representing the business intention toward environmental sustainability in the Amazon rainforest. (production/extraction means). From the hands of the figures/products seeking to contemplate this dimension, 5 out of the 11 items were adapted, dealing with the brand quality perception of the questionnaires developed by Netemeyer et al. (2004) and Yoo et al. (2017).

The third dimension contains three items, which seek to understand how many participants would be willing to buy products from companies related to environmental sustainability through the presented data (willingness to buy - WTB). To measure willingness to buy, the 3 (three) arranged items according to Arnett et al. (2003).

Finally, the fourth dimension assesses how much the business presents branding innovation (RI). The dimension was set up by asking participants to answer four questions that addressed the theme of innovation, uniqueness, and brand design, and the questions
were adapted from the questionnaires proposed by Lin, (2015) and Netemeyer et al. (2004). Of these works, only the translation of the items was carried out.

Seeking to comply with ethical principles in research with human beings, this study was submitted to the Research Ethics Committee, receiving the code CAAE: 55999722.1.0000.5546. Participants who agreed to participate in the study were informed about the research objectives, data collection, and the information regarding the Free and Informed Consent Term (FICT).

2.3 CROSS-CULTURAL TRANSLATION OF THE ITEMS INTO THE INSTRUMENT

The translation and adaptation of the items used from the original questionnaires followed five essential steps recommended by the literature: (1) translation of the instrument from the source language to the target language, (2) analysis of the prototypes and analysis of the synthesized version by expert judges, (4) back-translation to the source language, and (5) pilot study (Gjersing et al., 2010; Hambleton, 1993). For the convenience of this research, another task (phase 2) was performed during content validation by CVC for the full scale, and phases (4) and (5) were performed after the completion of content validation.

2.4 SCALE VALIDATION

After the formulation and structuring of the items to be used in the data collection scale, it was submitted for validation based on the tripartite model described in the Standards for Educational and Psychological Testing, (Frey and Association, 2018), The Standards present five validation processes to accept the instrument as reliable and replicable:

i. Content validation;
ii. Validity based on internal structure;
iii. Validity Based on Relationships with External Measures if
iv. Validity based on response pattern to items.
2.4.1 Content validity

The content validity coefficient (CVC) was applied in the proposed instrument to measure the effectiveness of the scale based on qualitative evaluations of expert judges in the area. This technique involves seeking the consensus of a community of professionals who will attribute values and opinions about the translations and adaptations made in the instrument so that it can be refined and contemplate the initial objectives of the instrument. The evaluation of the scale by the judges was performed blindly; therefore, the confidentiality of the participants guarantees greater fidelity to the results at the end of the experiment, allows better exposure to the results, and maintains their understanding (Hasson et al. (2000). After item validation, randomization was performed using Excel software, and data collection was carried out from March to May 2022.

2.4.2 Pilot public and initial application of the tool.

In order to cover a higher number of regions in the country, the instrument was applied both face-to-face and virtually. According to Walter and Block, (2016), due to its advantages, such as speed and capacity, the application of digital questionnaires has spiked to reach specific populations, reducing the possible costs of travelling. The selection of the participating respondents was randomly performed to signify population variability in Brazil.

2.5 PILOT VALIDATION OF THE TOOL

2.5.1 Validity is based on internal structure.

Validity based on internal structure reflects the degree to which the structure of correlations between items matches what the construct intends to measure (Frey & Association, 2018). Therefore, in this research, an AFE was carried out to evaluate the factor-named structure of the proposed scale. The analysis was implemented using a polychromic matrix, and the extraction method used total diagonally weighted least squares (RDWLS) (Asparouhov et al., 2010). The decision on the number of factors to
be retained was performed using the technique of parallel analysis with a random permutation of the observed data (Timmerman & Lorenzo-Seva, 2011).

To determine the number of retention factors and facilitate the interpretation of the factors (Damásio, 2012), a rotation method was adopted similar to the oblique one by Robust Promin (Ferrando & Lorenzo-Seva, 2018). Seeking a better refinement of the scale, it was assumed that, for an item to be accepted, it must have a factor loading above 0.60 and cross-loadings between items lower than 0.20 (Howard, 2016).

Seeking to evaluate the quality of the model's adequacy and to assess how well it can reproduce the covariance structure or the correlational structure of the variables, the following strategies were used: the root mean square error of approximation (RMSEA), Tucker-Lewis index (TLI) and comparative adjustment index (CFI). According to Brown, (2015), the RMSEA value must be less than 0.08, and the confidence interval cannot reach 0.10. The CFI and TLI values must be greater than 0.90, preferably 0.95. The model's reliability, that is, how consistent and reproducible it is in measuring the intended latent variables, was tested using the composite reliability indicator, which must have values greater than or equal to 0.70 (Hair et al., 2009).

The factor’s stability was evaluated using the H-index (Lorenzo-Seva & Ferrando, 2019). The H-index proposes to assess how well a set of items represents a common factor. H values range from 0 to 1. High H values (> 0.80) suggest a well-defined latent variable, which is likely to be stable in different studies.

2.5.2 Validity based on relationships with external measures

We sought to evaluate the evidence of the validity of the instrument based on the relationship with other variables arranged in other similar questionnaires (Souza et al., 2017), as well as the information asymmetry. For this, convergent and discriminant validity tests were performed. The application of convergent validity seeks to understand the significant relationship between two or more measures of the same construct or theoretically related constructs. (Cunningham et al., 2016). Discriminant validity is necessary as constructs to be theoretically valid should not be related to each other (Hubley, 2014). In this sense, the application of descriptive validation tests aims to
evaluate the difference between measurements of different constructs present in the instrument.

2.5.3 Convergent and discriminant validity
The factors convergent validity was evaluated utilizing the average variance extracted (AVE) and using standardized factor loading and the error of the item measure. The CVA estimation was obtained by the factor loadings and measurement errors obtained by confirmatory factor analysis using the *weighted estimator the least square mean and variance* (WLSMV). AVE values greater than 0.50 indicated good convergent validity (Maroco et al., 2014).

2.5.4 Validity of Invariance of Measures
Multigroup confirmatory factor analysis (MGCFA) was performed to verify the invariance of the proposed model measures in two samples selected at random to the extent of 50% for each (T=Test and A=Sample) of the collected data (J. C. Borsa & de Sousa, 2018; Callegaro Borsa et al., 2012). *Comparative Fit Index* (CFI) tests and the chi-square difference were used to assess the scale invariance. To assume the invariance of the measure, in comparison with the previous model, the fit index of the test model in the CFI must not be greater than 0.01 (Cheung & Rensvold, 2002), and the $\chi^2$ must present a value of $p > 0.05$ (Milfont & Fischer, 2010; Zaiț & Bertea, 2011).

2.5.5 Validity based on the item response pattern
Item response theory (IRT) techniques are sets of psychometric models used to describe the probability of an interviewee providing a specific and real answer in an instrument based on the parameters and perceptions of the item (latent characteristics), i.e., being able to identify the peculiarities of an interviewee that cannot be measured directly (de Souza Barbosa et al., 2021). The most used techniques, such as the Rasch mode, have the premise that for each response judgment, there is a unique latent trait in the research subject (Andrich, 1978).
The psychometric properties of the adapted instrument were evaluated using Rasch model for polychotomous data assuming two logistic parameters (difficulty and discrimination). In addition, indicators of the reliability of people and items and performance deviations were evaluated through the infit and outfit indices. Regarding reliability indices, values greater than 0.70 or, preferably, values greater than 0.80 are expected (Linacre, 2012). The Infit and Outfit indices quantify residuals for the items in the model tested (Bond & Fox, 2013; Linacre, 2012). The DIF was evaluated using the Mantel procedure (Wright et al., 1994). Items whose difficulty estimates were significantly different for men and women (p ≤ 0.05) were inspected. The magnitude of DIF was interpreted through DIF contrast: values between | 0.00 | and | 0.43 | are considered low/negligible; values between | 0.44 | and | 0.64 | are considered moderate; and values above | 0.64 | are considered high (Linacre, 2012).

3 RESULTS

3.1 CROSS-CULTURAL TRANSLATION OF THE ITEMS INTO THE INSTRUMENT

The process of translation and cultural adaptation generated the Brazilian Portuguese version. During the translation process, important difficulties were not found. However, the process has shown some idiomatic problems, resulting in some weak changes in grammar structure. Based on the translators' consensus, 23 adapted items were accepted, 19 polyatomic (Likert scale 1 - 5), two dichotomous (Yes - No), and one qualitative question were also accepted.

3.2 VALIDITY BY INTERNAL MEASURES

3.2.1 Content validity

Based on the scores attributed to the items, the CVC was calculated to evaluate the agreement that the judges perform on the scale. Based on the score assigned by the judges, items that did not reach a value greater than or equal to 0.8 were grouped and relocated for correction in at least two of the established criteria (clarity, practical relevance, and theoretical relevance).
When implementing the validation procedures under APA standards for cross-cultural adaptation in Portuguese-speaking population, the initial qualitative study, as expected, revealed initial translation overlaps, as well as important discrepancies in the languages used. From the base instrument in its original version, five changes were accepted in the statements and submitted to new rounds. RSA 02 “Can you quickly remember the brand symbol or logo?” was corrected to “Do you think this brand symbol/logo is easy to remember?” (clarity = 0.430); (relevance = 0.650); (relevance = 0.700); RI 01 “When imagining the product, I realized that the quality of the products is high” changed to “When imagining a product of this brand, I believe that the quality is high” (clarity = 0.896) (relevance = 0.963); (relevance = 0.963); QLM 01 “I feel that when I purchase the services, I will receive the quality I expect.” Changed to “I feel that by purchasing the services of this brand, the quality will meet my expectations.”, BQP02 (clarity = 0.830) (relevance = 0.963); (relevance = 0.896) “While viewing the services of the Brand, would you be able to differentiate of products from other companies?” changed to “After seeing the above image, would you be able to differentiate branded services from other companies?”; WTB 03 (clarity = 0.896) (relevance = 0.963); (relevance = 0.963) "If the Brand’s prices rise a little, I would buy from the competition, regardless the responsibility of environmental sustainability." changed to "If the brand’s prices went up a little, I would buy from the competition, even though I know they care about environmental sustainability." Other listed items to compose the scale met the established criteria.

The change in item RI 01 was suggested considering the change from the term "perceive" to "believe" since imagining that the item has some quality factor and is semantically closer to belief than to perception. Given that the CVC\textsubscript{t} presented satisfactory results for all aspects judged, the change seemed appropriate, proposing that the term was adaptable semantically.

After adjustments, the Content Validity Coefficient for the complete scale was accepted for eligibility of the instrument to the new validation phase (Santos, 2018). Considering the CVC\textsubscript{t} mean for three aspects judged — clarity, coherence, and relevance — was CVC\textsubscript{t} = 0.87. Based on the values found, CVC\textsubscript{t} for the entire instrument presented satisfactory values of language clarity (CVC\textsubscript{t} = 0.84), practical relevance (CVC\textsubscript{t} = 0.91)
and theoretical relevance ($CVC_t = 0.90$), having a content validity coefficient for the test as a whole of $CVC_t=0.87$ and an error $Pe_t=0.003$. The instrument was considered eligible to other phases, considering values similar to those found by Daffern et al. (2017).

3.3 VALIDITY BASED ON INTERNAL STRUCTURE

Bartlett's (2926.2, $gl = 210; p < 0.001$) and KMO (0.89) tests of sphericity suggested the interpretability of the items correlation matrix. For stability, Cronbach's alpha presented acceptable values (Cronbach's $= 0.925$). The parallel analysis proposed four more representative factors for the data (Table 1).

Table 1. Results of the parallel analysis of the scale of consumers' perception of sustainable brands on the Amazon rainforest

<table>
<thead>
<tr>
<th>Percentage of explained variance of actual data</th>
<th>Percentage of explained variance of random data (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>49.527 *</td>
<td>10.780</td>
</tr>
<tr>
<td>8.888 *</td>
<td>9.832</td>
</tr>
<tr>
<td>7.545 *</td>
<td>9.026</td>
</tr>
<tr>
<td>5.293 *</td>
<td>8.476</td>
</tr>
</tbody>
</table>

Note: The number of factors to retain is four, as four factors from actual data have a higher % explained variance than random data.

The loading factor of the items can be seen in Table 2. In addition, the composite reliability indices are reported, as well as estimates of replicability of the factor scores ($H$-index) (Ferrando & Lorenzo-Seva, 2018). From the factor analysis, it was possible to identify dimensions, that is, to reorganize the explanatory items in their best positioning, and to verify which ones did not reach the necessary factor loading.
Table 2. Factor structure of consumers’ perception of sustainable brands on the Amazon rainforest

<table>
<thead>
<tr>
<th>Items</th>
<th>SBP</th>
<th>BQP</th>
<th>RI</th>
<th>WTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP 01</td>
<td>0.451</td>
<td>0.041</td>
<td>0.244</td>
<td>-0.166</td>
</tr>
<tr>
<td>SBP 02</td>
<td>0.837</td>
<td>-0.240</td>
<td>0.222</td>
<td>0.061</td>
</tr>
<tr>
<td>SBP 03</td>
<td>0.921</td>
<td>-0.107</td>
<td>-0.054</td>
<td>0.059</td>
</tr>
<tr>
<td>SBP 04</td>
<td>0.899</td>
<td>-0.134</td>
<td>0.044</td>
<td>0.053</td>
</tr>
<tr>
<td>SBP 05</td>
<td>0.750</td>
<td>0.136</td>
<td>-0.149</td>
<td>0.132</td>
</tr>
<tr>
<td>SBP 06</td>
<td>0.823</td>
<td>0.141</td>
<td>-0.294</td>
<td>0.080</td>
</tr>
<tr>
<td>SBP 07</td>
<td>0.829</td>
<td>-0.045</td>
<td>0.212</td>
<td>-0.128</td>
</tr>
<tr>
<td>SBP 08</td>
<td>0.929</td>
<td>-0.007</td>
<td>0.094</td>
<td>-0.110</td>
</tr>
<tr>
<td>SBP 09</td>
<td>0.569</td>
<td>0.093</td>
<td>0.202</td>
<td>-0.245</td>
</tr>
<tr>
<td>BQP 01</td>
<td>-0.210</td>
<td>0.407</td>
<td>-0.083</td>
<td>-0.048</td>
</tr>
<tr>
<td>BQP 02</td>
<td>-0.270</td>
<td>0.706</td>
<td>-0.007</td>
<td>0.166</td>
</tr>
<tr>
<td>BQP 03</td>
<td>0.055</td>
<td>0.728</td>
<td>-0.041</td>
<td>0.064</td>
</tr>
<tr>
<td>BQP 04</td>
<td>-0.031</td>
<td>0.577</td>
<td>0.835</td>
<td>0.077</td>
</tr>
<tr>
<td>BQP 05</td>
<td>0.277</td>
<td>0.117</td>
<td>-0.013</td>
<td>-0.013</td>
</tr>
<tr>
<td>RI 02</td>
<td>-0.054</td>
<td>0.131</td>
<td>0.806</td>
<td>0.084</td>
</tr>
<tr>
<td>RI 03</td>
<td>-0.014</td>
<td>0.297</td>
<td>0.512</td>
<td>-0.125</td>
</tr>
<tr>
<td>RI 04</td>
<td>0.111</td>
<td>0.013</td>
<td>0.688</td>
<td>0.124</td>
</tr>
<tr>
<td>WTB 01</td>
<td>0.071</td>
<td>0.136</td>
<td>-0.047</td>
<td>0.464</td>
</tr>
<tr>
<td>WTB 02</td>
<td>0.133</td>
<td>0.263</td>
<td>-0.112</td>
<td>0.495</td>
</tr>
<tr>
<td>WTB 03</td>
<td>0.069</td>
<td>-0.045</td>
<td>0.013</td>
<td>0.813</td>
</tr>
<tr>
<td>WTB 04</td>
<td>-0.018</td>
<td>-0.021</td>
<td>0.207</td>
<td>0.833</td>
</tr>
<tr>
<td>WTB 05</td>
<td>0.251</td>
<td>0.041</td>
<td>0.244</td>
<td>0.464</td>
</tr>
</tbody>
</table>

Composite Reliability 0.937 0.703 0.714 0.760

H-latent 0.965 0.916 0.916 0.905

Note: Brand quality perception – BQP; Strategic brand positioning – SBP; Willingness to buy — WTB and Retail Innovation — RI.

The structure factor yielded few cross-loadings, and the loading factor was higher for initially expected factors than others. There was only one item (Retail Innovation) "The brand offers more innovative products than other companies I know" that did not project a fair value on the expected factor (i.e., items with factor loading above 0.30). Based on Nunnally (1978), items with loading factor below 0.30 were omitted. In others, only one item showed a pattern of cross loads (19); this item is allocated to the class with the highest load value and corroborates the findings in the literature. All items presented values close to the load of the dimension itself, also corroborating the discriminant validity, a behavior similar to that found by Castro et al., (2020).

The composite reliability of the factors was acceptable (above 0.70) for the SBP factor and marginally acceptable for the other. Second Ferrando and Lorenzo-Seva, (2018), the data suggest that all factors may be replicable in future studies ($H < 0.80$). In the composite reliability, the factors BQP, WTB, and RI presented values marginally
accepted by the literature. According to Valentini and Damásio, (2016), the change in the homogeneity of the factor loadings may be due to the smaller number of items in the factor to limit the interpretation of the results of the composite reliability.

The scale adjustment indices were adequate ($\chi^2 = 86,460$, $gl = 132$; $p < 0.001$; RMSEA = 0.027; BCa 95% [0.00 – 0.037]; CFI = 0.998; BCa 95% [0.998 – 1.002]; TLI = 0.997; BCa 95% [0.997 - 1.002]). In the Root Mean Square Error of Approximation (RMSEA), a value lower than 0.08 for the proposed instrument indicates a good statistical fit for the model (Tucker and Lewis, 1973; West et al. 2012). Comparative adjustment indices (CFI) and Tucker-Lewis index (TLI) above 0.95 also indicate a good instrument adjustment.

It is essential to highlight that the indicators of one-dimensionality Unidimensional Congruence (UniCo) 0.845, Explained Common Variance (ECV) 0.886, and Mean of Item Residual Absolute Loadings (MIREAL) 0.177 did not support the hypothesis of one-dimensionality of the scale. This is due to values of UniCo, ECV, and MIREAL being greater than 0.95, which suggests that the data can be treated as essentially unidimensional; otherwise, the null hypothesis is accepted, (Lorenzo-Seva & Ferrando, 2020), as for the MIREAL (Mean of Residual Item Absolute Loads) less than 0.300 suggests that the data can be treated as essentially one-dimensional. Therefore, this implies that our proposed instrument satisfies the benchmarks for the theoretical expression of formative constructs (Sarker et al., 2021).

Table 3 breaks down the quality and effectiveness of the factors. The factor determinacy index (FDI) can be interpreted as the number of different factor levels that can be differentiated based on the factor score estimates. Values above 0.90 are considered good for the FDI (DiStefano et al., 2018; Shi & Maydeu-Olivares, 2020).
Table 3. Quality and Effectiveness of Factor Score Estimates

<table>
<thead>
<tr>
<th>Methods</th>
<th>SBP</th>
<th>BQP</th>
<th>WTB</th>
<th>RI</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDI</td>
<td>0.974</td>
<td>0.951</td>
<td>0.950</td>
<td>0.948</td>
</tr>
<tr>
<td>ORION</td>
<td>0.951</td>
<td>0.904</td>
<td>0.902</td>
<td>0.878</td>
</tr>
</tbody>
</table>

Note: FDI = Factor Determinacy Index; ORION = Overall Reliability of fully Informative Prior Oblique N-EAP scores; Brand quality perception – BQP; Strategic Brand Positioning – SBP; Willingness to buy – WTB and Branding Innovation – RI.

The values obtained by the exploratory factor structure confirm the initial prediction proposed in this research for the construction of the instrument, establishing the final model with four factors. To brand quality perception factor (BQP) was assigned five questions describing the perception of the potential consumer interviewed when the evaluated brand presented the product. For the strategic brand positioning factor (SBP), ten questions were assigned. The SBP dimension is theoretically based on the concept of strategic positioning described by López and Alcañiz, (2000). The authors define the brand strategic positioning as their performance in the search for factors that make up for the success of the business, such as technological innovation and relation to the ethical concepts of the consumer and the environment. Finally, the willingness-to-buy factor (WTB) was composed of 2 questions, and the retail innovation factor (RI) was composed of 4 questions.

3.4 VALIDITY BASED ON RELATIONSHIPS WITH EXTERNAL MEASURES

3.4.1 Convergent and discriminant validity

The proposed model from the validity by internal measures comprises a total of 22 items divided into four dimensions, namely: SBP (0.845), BQP (0.964), WTB (0.889), and RI (0.677). According to Maroco et al. (2014), all dimensions reached acceptable values (> 0.50). The analyzed validity pairwise showed good values to attest validity (Table 4). According to Maroco et al. (2014), the accepted VEM value for each dimension must be greater than 0.50. Furthermore, the operational measures obtained by the factors, when analyzed pairwise, showed satisfactory values to attest to the validity hypothesis when analyzing the discriminant validity of the instrument.
Table 4. Discriminant validity of scale factors.

<table>
<thead>
<tr>
<th>Correlation of dimensions</th>
<th>$Z$ correlation coefficient</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP ↔ BQP</td>
<td>0.699</td>
<td>0.488</td>
</tr>
<tr>
<td>SBP ↔ WTB</td>
<td>0.846</td>
<td>0.715</td>
</tr>
<tr>
<td>SBP ↔ RI</td>
<td>0.857</td>
<td>0.734</td>
</tr>
<tr>
<td>BQP ↔ WTB</td>
<td>0.596</td>
<td>0.355</td>
</tr>
<tr>
<td>BQP ↔ RI</td>
<td>0.672</td>
<td>0.451</td>
</tr>
<tr>
<td>WTB ↔ RI</td>
<td>0.756</td>
<td>0.571</td>
</tr>
</tbody>
</table>

According to Fornell and Larcker (1981), discriminant validity is attested when one of two factors ($fej$) is greater than or equal to the square of the correlation between these factors. Then, the factors explain more variability of the items that are manifestations of each of the factors than the variability that only one factor explains (J. Borsa & Damasio, 2018). Reaching satisfactory indices, the instrument was considered eligible for the other validations.

3.4.2 Measurement invariance

Table 5 shows the configurable, metric, and scalar multigroup invariance analysis. These are the results of the scale in the different investigated groups. The data of different tests of the *Comparative Fit Index* (ΔCFI) indicate that the measure presents fine invariance adjustments. To assume measurement invariance, the value assigned to (ΔCFI) must not be greater than 0.01. It was observed, through the ΔCFI values, that the scale structure is stable, with no biased response in samples for the different groups (Lee & Skorski, 2019).
Table 5. Comparison of models by ΔCFI for the applied scale Multigroup Confirmatory Factor Analysis

<table>
<thead>
<tr>
<th>groups</th>
<th>Quiz</th>
<th>$X^2$</th>
<th>df</th>
<th>P Value</th>
<th>RMSEA (90% CI)</th>
<th>CFI</th>
<th>TLI</th>
<th>ΔCFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configurable invariance</td>
<td>2413.245</td>
<td>421</td>
<td>0.00</td>
<td></td>
<td>0.079 (0.062 0.086)</td>
<td>0.952</td>
<td>0.968</td>
<td>---</td>
</tr>
<tr>
<td>Metric invariance</td>
<td>1177.096</td>
<td>383</td>
<td>0.00</td>
<td></td>
<td>0.072 (0.065 0.079)</td>
<td>0.951</td>
<td>0.997</td>
<td>0.001</td>
</tr>
<tr>
<td>Scalar invariance</td>
<td>1186.748</td>
<td>400</td>
<td>0.00</td>
<td></td>
<td>0.065 (0.059 0.072)</td>
<td>0.953</td>
<td>0.926</td>
<td>-0.001</td>
</tr>
</tbody>
</table>

The invariance data by multigroup confirmatory factor analysis (AFCMG), through the values of configurable, metric, and scalar invariance, show that structures measured between the groups and the scalar units are the same for the groups studied since the values of difference from the Comparative Fit Index (ΔCFI) are below 0.01. Factor invariance is an essential component of the iterative process of demonstrating the measurement equivalence of latent constructs across groups, including gender and age subpopulations.

When measurement equivalence is present, the relationship between the latent variable and the observed variable remains unchanged across the analysis populations, (Dong & Dumas, 2020). This premise succinctly demonstrates the importance of measurement equivalence in psychometric instruments, suggesting that it is possible to apply this scale when the objective is to analyze, in different groups, the perceptions of consumers concerning brands related to environmental sustainability in the Amazon rainforest.

3.4.3 Validity based on the item response pattern

The validity based on the response pattern to items using Item response theory (IRT) when associated with invariance analysis allows the evaluation of the similarity of items in a given instrument developed for different groups (Sireci, 2021). The initial analyses showed adequate reliability indices for the items (Reliability = 0.98; Separation Index = 7.48) and for people (Reliability = 0.90; Separation Index = 2.93), which suggests that the estimates obtained tend to be replicable. Regarding the performance deviations (infit and outfit), it is noteworthy that the instrument presents acceptable values, (Bond & Fox, 2013; Boone, 2016), both for items (Infit MNSQ =1.00, ZSTD = -0.70; Outfit MNSQ
= 1.06; ZSTD = -0.13) and for participants (\textit{Infit} MNSQ = 1.10, ZSTD = -0.10; \textit{Outfit} MNSQ = 1.06; ZSTD = -0.20).

The results of IRT buttress the validity and reliability of the instrument. Furthermore, the values of reliability of the items (0.98) and the separation index of the items (7.48), as well as the index of the reliability of the participants = 0.90, and the index of separation of the participants (2.93), suggest that the estimates obtained tend to be replicable. \textit{Infit}s and \textit{outfit} values of the instrument of the items and the participants prove to be acceptable (Martinková et al., 2017).

The item-person map (Figure 1) shows that in the presented instrument, the level of latent trait was greater than the difficulty of items (mean theta = 0.67; mean difficulty of the items = 0). As seen, the most difficult items were BQP 02: “After seeing the image above, can you differentiate the services of this company from other companies?” (d = 1.83) and RI 01: "Do you believe that this brand symbol/logo is easy to remember?" (d = 1.43), and the most specific item was item SBP 07, “I would consume this mark” (d = -0.96).

Figure 1. Person-item map for each dimension of the validated scale.
The majority of items are grouped between – 1.5 and approximately 1.5 logi (Figure 1). So there is greater precision in the *theta estimation* of the participants allocated in this continuum range, and Test Information Curve grounds such information. Regarding the statistics of the items individually, Table 6 shows the threshold indicators when all items showed an increasing structure, as theoretically expected.

<table>
<thead>
<tr>
<th>Items</th>
<th>Thresholds 1-2</th>
<th>Thresholds 2-3</th>
<th>Thresholds 3-4</th>
<th>Thresholds 4-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP 01</td>
<td>-7.690</td>
<td>-2.420</td>
<td>-1.360</td>
<td>-0.400</td>
</tr>
<tr>
<td>SBP 02</td>
<td>-8.420</td>
<td>-3.150</td>
<td>-2.090</td>
<td>-1.130</td>
</tr>
<tr>
<td>SBP 03</td>
<td>-8.110</td>
<td>-2.840</td>
<td>-1.780</td>
<td>-0.810</td>
</tr>
<tr>
<td>SBP 04</td>
<td>-7.910</td>
<td>-2.640</td>
<td>-1.580</td>
<td>-0.620</td>
</tr>
<tr>
<td>SBP 05</td>
<td>-7.290</td>
<td>-2.020</td>
<td>-0.960</td>
<td>0.000</td>
</tr>
<tr>
<td>SBP 06</td>
<td>-7.700</td>
<td>-2.430</td>
<td>-1.370</td>
<td>-0.410</td>
</tr>
<tr>
<td>SBP 07</td>
<td>-8.520</td>
<td>-3.250</td>
<td>-2.190</td>
<td>-1.230</td>
</tr>
<tr>
<td>SBP 08</td>
<td>-8.210</td>
<td>-2.930</td>
<td>-1.870</td>
<td>-0.900</td>
</tr>
<tr>
<td>SBP 09</td>
<td>-8.000</td>
<td>-2.730</td>
<td>-1.670</td>
<td>-0.710</td>
</tr>
<tr>
<td>BQP 01</td>
<td>-5.740</td>
<td>-0.460</td>
<td>0.590</td>
<td>1.560</td>
</tr>
<tr>
<td>BQP 02</td>
<td>-7.280</td>
<td>-2.010</td>
<td>-0.950</td>
<td>0.100</td>
</tr>
<tr>
<td>BQP 03</td>
<td>-6.140</td>
<td>-0.870</td>
<td>0.190</td>
<td>1.160</td>
</tr>
<tr>
<td>BQP 04</td>
<td>-7.910</td>
<td>-2.630</td>
<td>-1.570</td>
<td>-0.610</td>
</tr>
<tr>
<td>BQP 05</td>
<td>-7.350</td>
<td>-2.080</td>
<td>-1.020</td>
<td>-0.060</td>
</tr>
<tr>
<td>RI 02</td>
<td>-7.320</td>
<td>-2.050</td>
<td>-0.990</td>
<td>-0.020</td>
</tr>
<tr>
<td>RI 03</td>
<td>-7.900</td>
<td>-2.630</td>
<td>-1.570</td>
<td>-0.600</td>
</tr>
<tr>
<td>RI 04</td>
<td>-8.020</td>
<td>-2.750</td>
<td>-1.690</td>
<td>-0.730</td>
</tr>
<tr>
<td>WTB 01</td>
<td>-6.550</td>
<td>-1.270</td>
<td>-0.210</td>
<td>0.750</td>
</tr>
<tr>
<td>WTB 02</td>
<td>-7.540</td>
<td>-2.270</td>
<td>-1.210</td>
<td>-0.250</td>
</tr>
<tr>
<td>WTB 03</td>
<td>-7.610</td>
<td>-2.340</td>
<td>-1.280</td>
<td>-0.320</td>
</tr>
<tr>
<td>WTB 04</td>
<td>-7.720</td>
<td>-2.450</td>
<td>-1.390</td>
<td>-0.430</td>
</tr>
<tr>
<td>WTB 05</td>
<td>-7.690</td>
<td>-2.420</td>
<td>-1.360</td>
<td>-0.400</td>
</tr>
</tbody>
</table>

Regarding the DIF for age groups, Table 7 presents the results found. As seen, only RI 04 ("Do you think this brand symbol/logo is easy to remember?") showed significant DIF with moderate effect size (DIF Contrast = 0.48). The items with greater difficulty (BQP 02 and RI 04) reflect that potential consumers need a greater latent trait
for positive confirmation. This result shows that, for consumers, the visual signs used by brands do not reflect the strategic positioning sought.

To better understand this difficulty in items BQP 02 ("After seeing the image above, would you be able to differentiate the services of this company from other companies?") and RI 04 ("Do you believe that this brand symbol/logo is easy to remember?"), the results functioning differential analysis (DIF) support the premise that this RI 04 item presents varying DIF values when q values are observed for subjects with different age groups (Callegaro Borsa et al., 2012). The literature suggests that short-term recognition losses in adults may be linked to the activity performed in situations where attention is divided (Craik, 2018) – a fact that may have affected the participants of the older age group during the questionnaire in a digital environment.

Table 7. Differential Functioning of the Item (DIF) of the proposed instrument.

<table>
<thead>
<tr>
<th>Items</th>
<th>mantel chi-square</th>
<th>Sig.</th>
<th>DIF Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP 01</td>
<td>0.058</td>
<td>0.808</td>
<td>0.26</td>
</tr>
<tr>
<td>SBP 02</td>
<td>0.090</td>
<td>0.763</td>
<td>0.00</td>
</tr>
<tr>
<td>SBP 03</td>
<td>2.000</td>
<td>0.157</td>
<td>-0.63</td>
</tr>
<tr>
<td>SBP 04</td>
<td>0.272</td>
<td>0.601</td>
<td>0.12</td>
</tr>
<tr>
<td>SBP 05</td>
<td>2.000</td>
<td>0.157</td>
<td>-0.19</td>
</tr>
<tr>
<td>SBP 06</td>
<td>0.2727</td>
<td>0.601</td>
<td>-0.25</td>
</tr>
<tr>
<td>SBP 07</td>
<td>1.000</td>
<td>0.317</td>
<td>1.00</td>
</tr>
<tr>
<td>SBP 08</td>
<td>0.024</td>
<td>0.875</td>
<td>-0.23</td>
</tr>
<tr>
<td>SBP 09</td>
<td>2.000</td>
<td>0.157</td>
<td>-0.55</td>
</tr>
<tr>
<td>BQP 01</td>
<td>0.960</td>
<td>0.327</td>
<td>0.23</td>
</tr>
<tr>
<td>BQP 02</td>
<td>1.058</td>
<td>0.303</td>
<td>-0.32</td>
</tr>
<tr>
<td>BQP 03</td>
<td>2.454</td>
<td>0.117</td>
<td>0.46</td>
</tr>
<tr>
<td>BQP 04</td>
<td>0.058</td>
<td>0.808</td>
<td>-0.28</td>
</tr>
<tr>
<td>BQP 05</td>
<td>0.615</td>
<td>0.432</td>
<td>-0.16</td>
</tr>
<tr>
<td>RI 02</td>
<td>0.510</td>
<td>0.475</td>
<td>0.08</td>
</tr>
<tr>
<td>RI 03</td>
<td>3.240</td>
<td>0.071</td>
<td>0.21</td>
</tr>
<tr>
<td>RI 04</td>
<td>4.840</td>
<td>0.027</td>
<td>-0.48</td>
</tr>
<tr>
<td>WTB 01</td>
<td>0.250</td>
<td>0.617</td>
<td>0.10</td>
</tr>
<tr>
<td>WTB 02</td>
<td>3.240</td>
<td>0.071</td>
<td>-0.10</td>
</tr>
<tr>
<td>WTB 03</td>
<td>0.620</td>
<td>0.430</td>
<td>-0.63</td>
</tr>
<tr>
<td>WTB 04</td>
<td>0.924</td>
<td>0.336</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Figure 2 illustrates the relationship between the instrument characteristic curve and the standard errors of the items in the instrument. The solid Instrument Information line (θ) represents the scale information function. The overall scale provided additional information in the range $0 \leq \theta \leq +4$, dropping to values close to 6. The instrument standard error (dotted red line) provides a visual reference on how the estimation precision varies in θ, with smaller values corresponding to better estimation accuracy (Bean & Bowen, 2021). For errors, the estimated score accuracy on the proposed instrument is ranged between $0 \leq \theta \leq +6$.

Another required item that required a higher level of θ for positive endorsement (agree and agree) was WTB01, which the consumer is willing to pay a higher price (premium) up to a 10% limit. The results are close to those obtained by D’Amico et al. (2016) when analyzing the probability that a consumer will pay (or not) a premium price for a product with sustainable certification and that the brand equity sustainability individually does not promote the consumer’s interest in paying more for the product. Corroborating the findings by Rahmani et al. (2019).
The information faculties for each item are best illustrated by the item characteristic curves (ICCs), as shown in Figure 3a. These curves show that item information is not a static quantity, but conditional on the item's difficulty levels. For example, as shown in the item-person map, item WTB 01 shows more significant progress toward the respondent's endorsement level (θ) for positive responses to the questionnaire (I agree and agree).

The study contributes to the branding literature by proposing and verifying a model that examines the associations of latent variables that form brand equity (brand quality perception, strategic brand positioning, willingness to buy, and retail innovation) in brands that develop sustainable activities. From a model thus suggested, it is possible to understand the nomological network that connects the aspects of branding and sustainability of ecologically responsible enterprises.
Figure 3b describes the characteristic curve of the CCE scale. The curve shows that, for the analyzed instrument, a score of $\theta = -1$ translates into an expected score of 30 points; whereas a value of $\theta = 3.4$ translates into an expected valid score of 90 points.

On the other hand, item SBP 07, which presented the lowest level for positive endorsement (I agree and agree), shows that the consumer after knowing all the information about the brand and agreeing with the value of the products, plus a maximum of 10% premium increase, shows a tendency to consume the brands presented in the survey. It is important to note that items SBP 07 ($\theta = -1.28$) and WTB 02 ($\theta = -1.18$) presented a low-level $\theta$, which signals positive endorsements on the Likert scale (agree and agree) and shows that, after facing all brand information added to the agreement with the prices of the products, this potential consumer a maximum of a 10% premium increase, showing desire to consume the brands presented in the survey.

The instrument information curve briefly explains how well items, in general, provide statistical information about the latent trait. The curves are mathematical functions of each other, where Standard Error ($\theta) = 1 - \sqrt{\text{Information (} \theta) \). Higher information along the $\theta$ scale leads to lower standard errors, resulting in more accurate $\theta$ estimates. In addition, the scale information values can be used to calculate which level of $\theta$ the instrument reliably fits best. The range of information obtained by the validated items being between $0 \leq \theta \leq +4$ indicates that the instrument best applies to potential consumers who tend to respond moderately (between disagree and agree).
The transformation of the endorsement estimates \( \theta \) present in the Characteristic Curve of the Instrument makes it possible to verify that the test participants, on average, responded 50% above the positive score of the scale \( \geq 4 \) (I agree on the Likert scale). The transformation to the original scale metric provides a more familiar frame of reference to interpret the sum of scores, (Bean & Bowen, 2021), having properties favorable to reliability in the analysis of the latent trait, improving and making the approach in more accurate scores (Edwards & Wirth, 2009). Regarding the effects of the ceiling and floor, no participants were identified or questions that corroborated this occurrence for this instrument.

The final scale with 22 items, available in electronic version, has good psychometric properties. Pilot questionnaires based on works by Netemeyer et al. (2004), Yoo et al., (2017), Grigorescu et al. (2019), and Lin, (2015) within a total of 32 items covering five dimensions were reduced to 20 items covering four dimensions. Each dimension contains items about the perception of product quality, the perceived positioning of the brand, the willingness of the potential consumer to buy, and the innovation perceived in the branding of the analyzed brand.

The process of developing psychometric instruments by cross-cultural adaptation is an approach that can be applied to different cultural and linguistic environments (Cozendey-Silva et al., 2016). The inclusion of Portuguese-speaking countries can help fill a data gap on the nomological network to represent consumer perception of sustainable brand equity. The objective of adaptation is such that the primary goal of this instrument is to achieve equivalence between the instruments of the original measurements and their adapted versions (Beaton et al., 2007). In that sense, the proposal version for the Portuguese language of this instrument was obtained through careful steps for validation of instruments described in the recommendations for APA test standards (Frey & Association, 2018).

In addition to the instrument validation data, the initial results suggest that among the dimensions of sustainable brands, they exert a positive influence on the variable result of brand value for the consumer when it is intended to add premium value to the products.
The Portuguese version of the proposal has the potential to generate indicators and feedback, as the study's findings are in line with the results of previous authors, (Ng et al., 2014; Shanti, 2016) suggesting that this scale is capable of assessing data similar to those proposed in previous scales in English. However, an investigation of the resulting data through experimental comparison research should be carried out in the future, correcting it, measuring it, and making this instrument more robust (Stavropoulos et al., 2019).

Although this research targets at developing and validating a psychometric scale, the present findings contribute to the scientific findings that study sustainable brands, as effective branding plans require broad knowledge about the perception of consumers and potential consumers regarding the dimensions responsible for the strategic positioning of brand equity.

4 CONCLUSION

The proposed scale was psychometrically solid, presenting expected relationships with the known variables when associated to the perception of potential consumers regarding validity and reliability by standards. Therefore, the validated instrument can be an award-winning research tool in the observation of the perception of Brazilian consumers in different groups of their population and in the evaluation of branding management for future sustainable brands, since the evaluation of constructs that shape consumer perception is a complex phenomenon.

This scale has good psychometric properties and is very quick to complete. It can be easily used by entrepreneurs when the objective is to understand their brand in the face of consumer perception. This study yields several recommendations related to sustainable brand research. These include the positive perception of sustainable products, the acceptance of paying a premium price of up to 10% for products from sustainable brands, and the differentiation of logo memorization for different ages.

This scale is advantageous for company managers to boost their business in a sustainable and continuous way, as well as brand managers aimed at consumers of forest derived products. It is recommended that future studies should investigate further
evidence of validity, including additional measures of cognitive functioning, modeling by structural equations, and qualitative comparative analysis (QCA) with qualitative data from stakeholders involved in the brands, applying them in group exams to further test its usefulness.

Therefore, the hope is that this study serves as a starting point for future research to build and expand our knowledge on this scientifically interesting and management-relevant topic.

POTENTIAL COMPETING INTERESTS

The authors declare no competing aspirations.

DATA AVAILABILITY STATEMENT

All data generated or analyzed during this study is included in this published article (and its supplemental information files).
REFERENCES


Daffern, T., Mackenzie, N. M., & Hemmings, B. (2017). Predictors of writing success: How important are spelling, grammar and punctuation?:


The moderating effects of game-playing time and living with parents. Addictive Behaviors Reports, 9, 100137. https://doi.org/10.1016/J.ABREP.2018.10.001


