ESTUDOS EM DESIGN
DESIGN ARTICLES

Design orientado para a emoção: Validação com adaptação transcultural do PrEmo®

Emotion-oriented design: Validation with cross-cultural adaptation of the PrEmo® tool

Ana Laura Alves, School of Architecture, Arts, Communication and Design, São Paulo State University, Brazil
ana.laura-alves@unesp.br

Gerusa Ferreira Lourenço, Occupational therapy department, Federal University of São Carlos, Brazil
gerusa@ufscar.br

Luis Carlos Paschoarelli, School of Architecture, Arts, Communication and Design, São Paulo State University, Brazil
luis.paschoarelli@unesp.br

Resumo
Designers têm aplicado conhecimento orientado por emoção no processo de design de produtos e serviços usando métodos e ferramentas para coletar dados e orientar sua interpretação. Contudo, faz-se necessário validá-los para evitar interpretações errôneas e insights tendenciosos. Este estudo teve como objetivo conduzir o primeiro processo de validação do PrEmo® (versão estática) no Brasil. A amostra incluiu 692 participantes (18 - 80 anos). Os resultados mostraram que a taxa de reconhecimento para valência pretendida obteve bons resultados, mas não atingiu um nível confiável em relação à emoção pretendida. Com base nos resultados, identificou-se quais personagens do PrEmo® precisam de aperfeiçoamento, bem como definiu-se algumas diretrizes ou estratégias para interpretar os dados emocionais coletados pelo instrumento com mais precisão.

Palavras-chave: Emoção, Mensuração emocional humana, Instrumento de autorrelato não verbal, Validação transcultural, Design Emocional

Abstract
Designers have applied emotion-driven knowledge in the product and service design process using methods and tools to gather data and guide its interpretation. However, validating them is necessary to avoid data misinterpretation and biased insights. Therefore, this study aimed to conduct the first validation process of PrEmo® (card-set version) in Brazil. The sample included 692 participants (18 - 80 years old). The results showed that the recognition rate for intended valence had good results but did not achieve a reliable level regarding intended emotion. Based on our results, it was possible to identify which PrEmo® characters required improvement and to define some guidelines or strategies to interpret the emotional data collected by the instrument more accurately.

Keywords: *Emotion, Human emotional measurement, Non-verbal self-report tool, Cross-cultural validation, Emotional design*

Introduction

Over the centuries, the concept of emotion evolved as far as theories were proposed by researchers from the biology, psychology, and neuroscience fields, such as Darwin (2009), Plutchik (2002), Damasio (2021), and Porges (2011). These theories have stimulated the creation and enhancement of methods and measurement instruments for human emotional responses, which have seen significant growth over the last two decades (Cardello; Jaeger, 2021). Emotion is a complex phenomenon that covers neuro-physiological processes, behavioral expressions, and subjective perception. Each part of this human experience demands specific methods and tools (Scherer, 2005; Desmet, 2018). It is possible to measure human emotional responses through instruments with a physiological, behavioral, or self-report basis.

Regarding the instruments with a self-report basis, the literature presents verbal and non-verbal self-report instruments, for example, the Geneva Emotion Wheel (Bänziger; Tran; Scherer, 2005), Affect button (Broekens; Brinkman, 2013), Product Emotion Measurement Instrument (PrEmo®) (Desmet, 2019), and emoji scales (Santos *et al*., 2022). For Scherer (2005), researchers can apply these instruments during interviews or on questionnaires to gather the participants' perception, considering the whole emotional experience as a subjective cognitive representation. In the Design field, it is possible to find recent studies delving into emotion as a result of the researchers and designers' yearning to control the users' emotional responses to their designs (Desmet, 2018; Damazio; Tonetto, 2022) and its potential for design innovation (Peng; Desmet; Xue, 2023), as can be seen in Paschoarelli, Moreira and Silva (2024) who applied PrEmo®. This Dutch tool was created to address issues faced by designers and other human-centered practitioners during product development and user evaluation. Nonetheless, validated self-report instruments are not always adopted, as described by Alves *et al.* (2023). This methodological choice can influence data understanding and the comparison of results with those of previous studies.

The linguistic translation (Ozolis *et al*., 2020) with cross-cultural adaptation is pointed out as essential to safeguard the instrument's validity and reliability (Beaton *et al*., 2000). These authors mentioned validation for self-report instruments built up by words. However, studies (Laurans; Desmet, 2017; Jaeger; Ares, 2017; Desmet, 2018) also indicated the cultural influence on the pictorial instruments' perception and meaning judgment. According to Hofstede, Hofstede and Minkov (2010), culture gathers human behavior patterns grounded on habits and social rules related to how people think, feel, and act. Therefore, paying attention to the language and cultural validation of the self-report instrument to be adopted in a study for collecting emotional data is relevant. Some previous studies developed in Brazil (Gobbi *et al*., 2017; Duarte *et al*., 2022) applied PrEmo® without this in mind.

Based on the current literature on measure-emotion methods and the challenges of using self-report instruments, this study aimed to conduct a validation process with a cross-cultural adaptation of PrEmo® card-set version in Brazil. Laurans and Desmet (2017) indicated that running validation studies with PrEmo® in other countries and cultures may represent an advantage for designers and other practitioners. We did not find studies that described the validation of the current PrEmo® (Desmet, 2019) with static characters. Besides that, we aimed to compare our results with those of Laurans and Desmet (2017) to verify inconsistencies. The use of Premo® in research to gather emotional data from Brazilians without validation may raise

inquiries regarding the data's reliability and interpretation. Therefore, this study may provide theoretical and practical support to Brazilian designers and other practitioners who perform emotional evaluations of products and services, contributing to emotion-driven knowledge with practical and theoretical implications, as well as to Emotion measurement in the Design field.

Research Methodology

We performed an online survey without researcher moderation. The participants answered a questionnaire using a computer or smartphone. The validation process was divided into two phases due to language differences. The first phase aimed to translate the names of each emotion included in the present study into a coherent Portuguese version. The second phase focused on the semantic judgment survey in which participants matched PrEmo® characters with emotions. The Ethics Committee approved this study (CAAE 52909821.9.0000.5663).

Phase 1: Judges

The judges validated the proposed translations in this phase, minimizing possible biases. The judges evaluated the accuracy and coherence of the translations of terms extracted from previous studies (Scherer, 2005; Laurans; Desmet, 2017; Santos et al., 2022; Fokkinga; Desmet, 2022) that express emotions from English to Portuguese, considering Brazilian culture and everyday expressions. This process was adopted for two reasons. First, machine learning tools still have some limitations regarding cross-cultural adaptation to the Brazilian Portuguese language, and therefore, technical support from judges was necessary. Second, the process of this phase of the study is also practiced in translation methodologies and idiomatic validation of self-report instruments (Mokkink *et al*., 2019; Ozolis *et al*., 2020).

It is essential to highlight the importance of this phase, as its results were used to create the questionnaire used in Phase 2 of the current study and will be a guide for interpreting the results of research that may apply PrEmo® to collect emotional data.

Sample and Sampling

Researchers and practitioners participated as judges, all native Portuguese speakers with English fluency (n = 14). Three of them were English teachers, five researchers with master's degrees in Design, and six lecturers with PhD degrees in Psychology (2), Design (3), and Occupational Therapy (1). Except for the English teachers, the other judges had deep experience with product and emotional evaluation in Design. Recruitment followed judgmental sampling.

Stimuli

The terms' selection was based on the findings of Scherer (2005), Laurans and Desmet (2017), Santos *et al*. (2022), and the Emotion Typology (Fokkinga; Desmet, 2022). We were not limited to the terms incorporated in the final version of PrEmo® because this could bias the participants' semantic judgment. The judges evaluated the translation of 31 terms mainly associated with emotions. The terms applied were: Positive emotions – 'Admiration', 'Affection', 'Calmness', 'Desire', 'Excitement', 'Fascination', 'Hope', 'Joy', 'Love', 'Pleasure', 'Positive surprise', 'Pride', 'Satisfaction'; Neutral emotion – 'Neutral'; Negative emotions – 'Anger', 'Boredom',

‘Confusion’, ‘Contempt’, ‘Disappointment’, ‘Disgust’, ‘Distress’, ‘Distrust’, ‘Embarrassment’, ‘Fear’, ‘Frustration’, ‘Nervousness’, ‘Sadness’, ‘Shame’, ‘Shock’, ‘Tiredness’, and ‘Worry’.

Procedure and Data Analysis

The judges received a form containing research information and instructions on what and how they would evaluate the translations of the terms proposed by the authors. Based on the term definition and its context of use, the judges classified each translation as Adequate, Doubtful, or Inadequate. If they did not agree with the proposed translation and/or had any suggestions, there was space for justification. There was no image of PrEmo® on the form, so the characters’ facial expressions would not influence the judge’s assessment of the translation/word. The semantic judgment of the tool should only be carried out by Phase 2 participants.

During this evaluation process, the judges operated independently, with no interference from the authors in their review, except to clarify any questions. An example of how to fill out the form is shown in Figure 1. The full form can be seen in Appendix A.

Figure 1: An example of a judge review in the form supplied by the authors

	Term in English	**Translation proposed**	**Adequate**	**Doubtful**	**Inadequate**	*Justification*
Affection	Definition: The feeling when you think about or interact with a person that you really like. You have the urge to be close and spend time with that person. Context: *“I have a great affection for these sneakers.”*	Adoração		X		I suggest other terms such as ‘Afeição’ or ‘estima’

Source: Authors (2025)

An inter-judge agreement analysis was performed using descriptive statistics, absolute and relative frequencies, to compare the proposed translation and the judges’ evaluations for each term. We defined a minimum rate of 75% of agreement between the judges’ translations and ours.

Phase 2: Semantic judgment survey

This phase aimed to understand how native citizens of Brazil interpret and associate the terms for emotions (n = 33) of Phase 1 and PrEmo® characters (n = 14) through a semantic judgment survey. We first verified the recognition rate for PrEmo® and the intended emotions for each character and thus compared it to the results from Laurans and Desmet (2017). The results of our semantic judgment survey may suggest important implications for the accuracy of PrEmo® in Brazil.

Sample and Sampling

This study was carried out throughout Brazil. The demographic census (Brasil, 2022) was the reference for the population size (203,080,756 inhabitants) and was used to define the minimum percentage participation of each large federal region (North - 8.3%, Northeast - 27.8%, Center - West - 7.4%, South - 14.4%, and Southeast - 42.1%) in the sample by quota sampling. Appendix

B provides a sample overview, including demographic and socio-economic characteristics. The eligibility criteria followed were: 1) age between 18 and 80 years old, 2) native Portuguese speakers, born and living in Brazil, and 3) should have at least a high school academic background. We considered only Brazilian residents due to their cultural and idiomatic specificities. Regarding the educational level, this decision was made based on the findings of Youngstrom and Green (2003) regarding the influence of education on emotional expression.

A total of 692 volunteers with a mean age of 34.85 years (SD ± 12.37) participated in this research phase. The sample size was established at a minimum of 385 participants, adopting an error margin of 5% and a confidence level of 95% (z = 1.96). We used the formula below, where *N* represents the population size, *e* the margin of error (percentage in decimal format), and z denotes the z-score corresponding to the confidence level.

$$\text{Sample size} = \frac{\frac{z^2 \cdot p(1-p)}{e^2}}{1+\left(\frac{z^2 \cdot p(1-p)}{e^2 \cdot N}\right)}$$

Stimuli

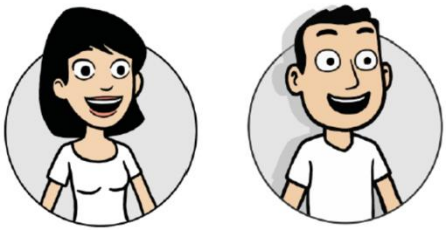
The stimuli conditions used in the questionnaire included the 14 characters of PrEmo® (independent variable) (Figure 2), representing seven positive and seven negative emotions by static facial and body expressions (Laurans; Desmet, 2017). Each emotion is represented by a character from the male and female genders (Desmet, 2019). PrEmo® promises to measure distinct emotions and mixed emotions. Thus, people can report their feelings by pointing out one or more characters to express their emotional experiences.

Figure 2: Non-verbal self-report PrEmo® tool

Source: Desmet (2019)

We used the card-set version without body movement and the vocal intonation created for the animated version (digital platform and app). In the questionnaire, female and male characters were positioned side by side to avoid potential embarrassment regarding gender identification (Figure 3). Besides that, the questionnaire contained the translated version of the 31 terms and two alternative options (‘no alternative’, ‘I do not know’) (dependent variable) (n = 33) to reduce cognitive confirmation bias. The translations can be seen in Appendix C.

Figure 3: Demonstration of how the characters were arranged in the questionnaire



Source: Desmet (2019)

Procedure and Data Analysis

We conducted a pilot test with five participants to identify potential shortcomings in the procedure and uncover opportunities for enhancing the questionnaire. It is important to note that the data from this pilot test was not included in the final sample.

The survey was carried out in Portuguese, and the participants answered a questionnaire with a fixed-choice judgment design. The questionnaire layout allowed participants to complete the questionnaire using a computer or smartphone without compromising the legibility and readability of the text or the visualization of the images. Each character was presented individually and judged through three questions. In the first question (*"In your opinion, which emotion represents better the figure above?"*), the characters were categorized with terms (n = 33). In the second question (*"In your opinion, this figure represents an emotion:")*, the participants judged the valence perceived of the characters considering four categories: 'positive', 'neutral', 'negative', and 'I do not know'. Lastly, in the third question (*"Have you already felt sometimes this emotion for a product? Understand as product physical objects (e.g., car, furniture, household utensils, toy, etc.) and digital (mobile application, website, virtual reality, etc.)"*) the participants indicated if they had a previous experience related to the emotion associated with that character when interacting with a product. This question also had four options: 'Yes', 'No', 'Maybe', and 'I do not remember'. Questions 2 and 3 were relevant to confirm the characters' emotional valence and the relevance of the emotion reported in Question 1 in the user experience evaluation. We randomized the presentation order of the stimuli conditions. The data were analyzed by descriptive statistics, using absolute and relative frequencies, following the analysis done by Laurans and Desmet (2017).

Results

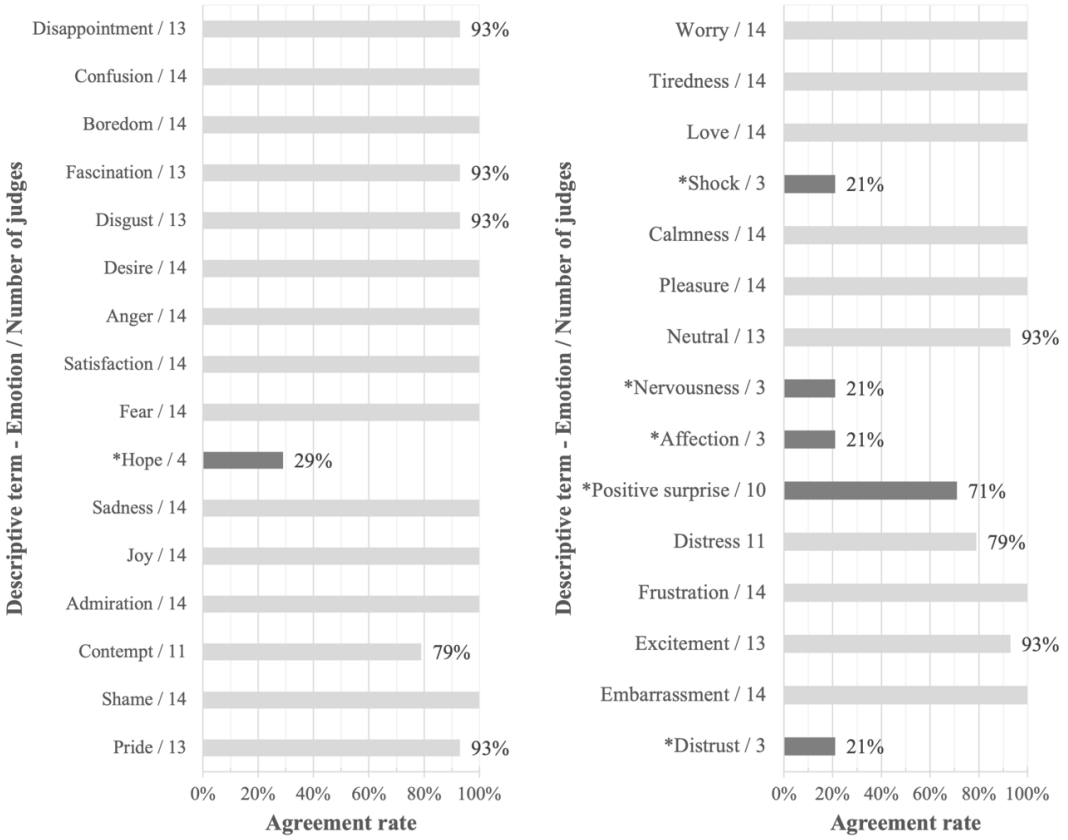
The results showed that most parts of the translations proposed for the terms had satisfactory agreement rates. Furthermore, the recognition rate for intended valence had good results for all characters but didn't achieve a reliable level regarding intended emotion.

Phase 1: Judges

After analyzing the inter-judge agreement, a total of twenty-five terms had agreement rates higher than 75%, with seventeen achieving 100% agreement (Figure 4). However, we noticed that six terms require revision to achieve a more accurate translation and understanding, without compromising the other research phases. Observing Figure 4, 'Shock', 'Nervousness',

'Affection', and 'Distrust' had an agreement rate of 21%. With better agreement rates than these terms, we found 'Hope' with 29% and 'Positive surprise' with 71%, but still lower than the ideal agreement rate of 75% (minimum). All translation changes followed the terms suggested by the majority of judges.

Figure 4: Results of the inter-judge agreement analysis



* Note: columns with darker gray indicate the terms that had translation into Portuguese changed (agreement rate < 75%).
Source: Authors (2025)

Phase 2: Semantic judgment survey

The recognition rate for intended emotion and valence for all PrEmo® characters are shown in Figure 5. The overall recognition rate for intended emotion was 46.29% (*SD* ± 23.46%), and the valence rate was 86,64% (*SD* ± 12.28%). The results suggested a satisfying valence rate, except the characters expressing 'Hope' were perceived as positive by 55% and as neutral by 34% of the participants. On the other hand, the recognition rate for intended emotion didn't achieve the expected results for a reliable non-verbal self-report tool for measuring emotion, with a hit rate below 50%.

When analyzing the recognition rates individually, the worst rates identified were 'Admiration' (12%), 'Fascination' (18%), 'Joy' (21%), 'Desire', 'Pride', and 'Shame' (33%). Among these characters, five express emotions with positive valence. Therefore, only two characters expressing positive emotions had satisfactory hit rates, 'Satisfaction' (48%) and 'Hope' (52%). Deepening the analysis, it's possible to observe that other terms had a higher hit rate than

the intended emotion. According to the participants' judgment, 'Joy' and 'Admiration' were associated with 'Positive surprise' with hit rates of 41% and 53%, respectively. 'Fascination' was more associated with 'Excitement' (25%). In the case of 'Pride', 'Satisfaction' had a hit rate of 35%, which was mixed with the character's intended emotion. There's a mixture of emotions. This confusion also occurred when participants judged the character expressing 'Satisfaction', but with a lower strength ('Pride' – 17%).

Figure 5: Results in percentage of the semantic judgment survey (n = 692)

Character – Intended emotion	Hit rate	Other choices above 10%	Valence	Other choices above 10%
Desire	33%	Affection (19%), Excitement (17%)	Positive (91%)	N[b]
Satisfaction	48%	Pride (17%) Calmness (10%)	Positive (85%)	Neutral (11%)
Pride	33%	Satisfaction (35%)[a]	Positive (79%)	Neutral (17%)
Hope	52%	Desire (10%)	Positive (55%)	Neutral (34%)
Joy	21%	Positive surprise (41%)[a] Excitement (14%) Admiration (11%)	Positive (98%)	N[b]
Fascination	18%	Excitement (25%)[a] Positive surprise (19%)[a] Admiration (11%)	Positive (82%)	Neutral (15%)
Admiration	12%	Positive surprise (53%)[a]	Positive (77%)	Neutral (15%)
Disgust	92%	N[b]	Negative (99%)	N[b]
Anger	48%	Disappointment (16%) Distrust (14%)	Negative (98%)	N[b]
Shame	33%	Embarrassment (24%) Disappointment (14%) Frustration (14%)	Negative (87%)	Neutral (10%)
Fear	57%	Shock (19%)	Negative (93%)	N[b]
Sadness	82%	Disappointment (10%)	Negative (98%)	N[b]
Boredom	51%	Tiredness (42%)	Negative (76%)	Neutral (23%)
Contempt	68%	N[b]	Negative (95%)	N[b]

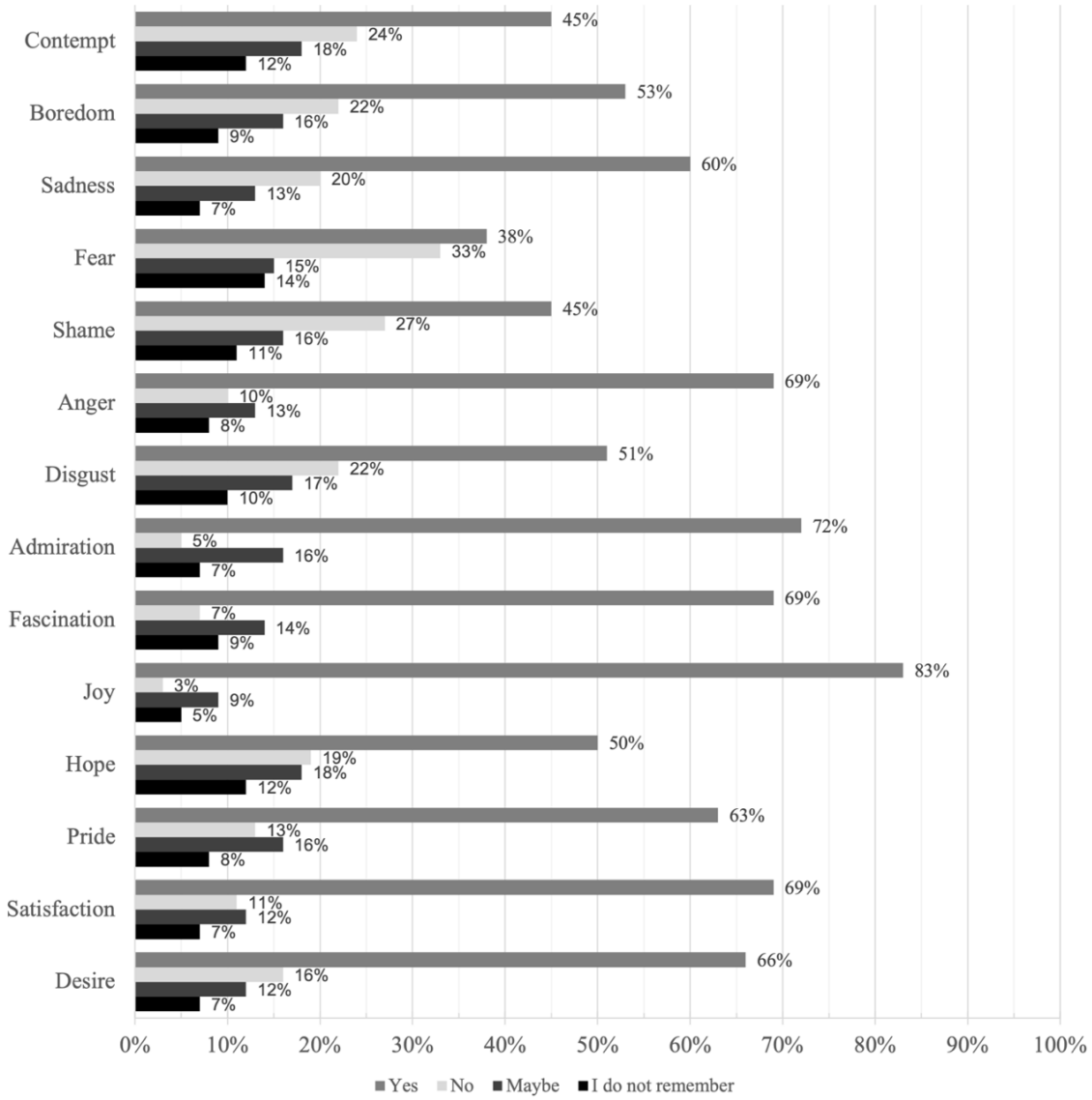
[a] Other terms with a higher hit rate than the intended emotion. [b] No other choices.
Source: Authors (2025)

The characters expressing negative emotions had higher hit rates than those expressing positive emotions, except for 'Shame', which had a 33% recognition rate for the intended emotion, and was also associated with 'Embarrassment' (24%), 'Disappointment' (14%), and 'Frustration' (14%). Another character with a negative valence that had interesting results was 'Boredom'. 'Boredom' had a hit rate of 51% for intended emotion, but the other choice with

notable association was ‘Tiredness’ (42%). These results showed some uncertainty in the participants’ judgment.

The general results concerning participants’ previous experience with the emotion associated with the characters during a product interaction are shown in Figure 6. Figure 6 synthesizes all data gathered without distinguishing which emotion the participant’s answer was associated with, the intended emotion, or the other choices above 10%. For example, 60% of the participants have already experienced ‘Sadness’ when interacting with products, but this result is related to the two terms with accuracy rates above 10%, ‘Sadness’ (82%) and ‘Disappointment’ (10%). Figure 6 shows the corresponding answers for ‘Yes’, ‘No’, ‘Maybe’, and ‘I do not remember’ for all emotions. These results identified the most evoked emotions in human-product interactions in agreement with Brazilian perception, such as ‘Joy’ (83%) and ‘Anger’ (69%).

Figure 6: Results in percentage of participants’ previous experience with the emotion associated with the characters during a product interaction



Source: Authors (2025)

Figure 7 presents the percentage of each emotion associated with the characters, considering both the intended emotions and other emotions chosen (> 10%).

Figure 7: Results in percentage of participants' previous experience with the emotion associated with the characters during a product interaction, considering intended emotion and other choices separately (n = 692)

Character - Intended emotion	General	Intended emotion and other choices above 10%	% For each emotion associated
Desire	66%	Desire Affection Excitement	86% 45% 76%
Satisfaction	69%	Satisfaction Pride Calmness	88% 53% 45%
Pride	63%	Pride Satisfaction	51% 85%
Hope	50%	Hope Desire	51% 78%
Joy	83%	Joy Positive surprise Excitement Admiration	82% 81% 86% 83%
Fascination	69%	Fascination Excitement Positive surprise Admiration	75% 80% 77% 76%
Admiration	72%	Admiration Positive surprise	85% 78%
Disgust	51%	Na	-
Anger	69%	Anger Disappointment Distrust	78% 67% 57%
Shame	45%	Shame Embarrassment Disappointment Frustration	29% 37% 68% 74%
Fear	38%	Fear Shock	33% 39%
Sadness	60%	Sadness Disappointment	58% 84%
Boredom	53%	Boredom Tiredness	62% 42%
Contempt	45%	Na	-

Source: Authors (2025)

We calculated the percentage of each term classified as other choices from their absolute frequencies. For example, 'Sadness'. A total of 566 participants (82%) associated the character

with the intended emotion, of whom 58% indicated that they felt this emotion in a product interaction, 21% didn't feel it, 14% maybe, and 7% didn't remember if they felt it before. All percentages in Figure 7 respect this calculation rule. Thus, the percentages changed when we distinguished the different terms associated with them, in some cases higher, in others lower.

On the one hand, the results demonstrated a fragile connection – rates lower than 50% – between the participant's emotional experience with a product and 'Shame' (29%), 'Fear' (33%), and 'Contempt' (45%), even considering the overall percentages of 45%, 38%, and 45%, respectively. With better results but still not ideal, we found 'Pride', 'Hope', and 'Disgust', with just 51% of the participants. On the other hand, when analyzing the rates of all terms associated with those characters that had problems regarding recognition rate for intended emotion ('Pride', 'Joy', 'Fascination', and Admiration'), we identified higher and similar rates of emotional experience with a product, except for 'Pride' (51%). So, the percentage of participants who indicated that they had felt those emotions before was (% intended emotion – % other choice): 85% 'Admiration' – 78% 'Positive surprise'; 75% 'Fascination' – 80% 'Excitement'; 82% 'Joy' – 81% 'Positive surprise'; 51% 'Pride' – 85% 'Satisfaction'.

Discussion

The current study aimed to undertake a validation process of the PrEmo® card-set version in Brazil. Based on the Brazilian sample's perception and judgment, we analyzed the recognition rate for the intended emotion and valence of PrEmo® characters and the participants' previous experience with the emotion associated with the characters during a product interaction. This work reports the entire validation process, including the cross-cultural adaptation of the terms used in PrEmo®, identifies interpretation conflicts, and brings new reflections on its reliability.

Our results showed that only six translations of terms needed revision, and all translation changes were made in line with the judges' suggestions. Despite not following a conventional translation process applied to the self-report measure method, our results indicated a great fit, a coherent and accurate adaptation of the terms attributed to emotions.

Regarding the recognition rate for intended emotion, the data analysis revealed a great difference between our results and those of Laurans and Desmet (2017). The recognition rate for all intended emotions, except for 'Disgust', had lower hit rates in this study compared to the previous study, especially the emotions with positive valence. For example, in the current study, 'Admiration' achieved a rate of 12%, which is far lower than the 85% reported by Laurans and Desmet (2017). This difference signals misinterpretations of the characters' facial expressions and, consequently, of the data collected with PrEmo®, depending on the culture. In other words, there may be a conflict between a participant's response in a product test and the interpretation of the data by researchers using PrEmo®. Laurans and Desmet (2017) said that the characters who expressed 'Joy' and 'Contempt' still had association problems (< 55%) with the intended emotion in the last validation test of the tool. Nevertheless, we found many more fragilities in the matching on PrEmo®.

Beyond the low recognition rates, the results indicated four positive emotions with dubious interpretations, 'Pride' (33%; Satisfaction – 35%), 'Joy' (21%; Positive surprise – 41%), 'Fascination' (18%; Excitement – 25%), and 'Admiration' (12%; Positive surprise – 53%). When

analyzing the hit rates of the intended emotion and the other choices (above 10%), we observed that the non-intended emotions had higher or proxy hit rates compared to those intended, which suggested that the participants had difficulty in recognizing the nuances of emotional expressions, a phenomenon described by Yoon *et al*. (2020) as Positive Emotional Granularity (PEG). According to Laurans and Desmet (2017), the inability to associate characters with their intended emotions may be related to the lack of clarity in the representation of emotional cues, differences in emotional granularity, and a few other contextual factors. Our findings suggest that culture may also be variable in the recognition rate, contradicting the assumption that characters are universally self-explanatory and cross-culturally consistent (Desmet, 2019). Feel different emotions are a natural human ability associated with the psychological (Hofstede; Hofstede; Minkov, 2010) and physiological functioning of human beings (Damasio, 2021). However, people learn how to deal with emotions and express them through culture, understood as the mental programming level shared among people who live or have lived within the same social environment, with their language, beliefs, and other common characteristics (Hofstede; Hofstede; Minkov, 2010). Therefore, our results suggest that the reliability of PrEmo® may fluctuate due to the cultural background of the users, such as participants in usability tests from Brazil, China, or the Netherlands. Despite Laurans and Desmet (2017) conducting tests with the animated version, Desmet (2019) stated that there was no difference in the tool's reliability depending on the version. However, the scientific findings of this study have raised doubts about the reliability of PrEmo®.

In respect of the recognition for intended valence, the results were satisfactory, indicating that the participants understood this emotional information. These findings corroborate the propositions of Lauran and Desmet (2017) and Desmet (2018; 2019). The only term that presented an unexpected hit rate was 'Hope', which had 55% for positive and 34% for neutral valence. This result can be associated with an issue of clarity in representation (Laurans; Desmet, 2017) and cultural bias, as discussed above.

Lastly, the results demonstrated that there is no straight relation between the recognition rate for intended emotion and the participant's experience with emotions evoked by a product. It is worth remembering that data on previous emotional experience (memory) associated with products was collected because there was a possibility that low recognition hit rates occurred due to participants not having experienced those emotions (at low rates) in interactions with products in their daily lives. Thus, the absence of repertoire may be an influencing variable in the semantic judgment of PrEmo®, which was not confirmed. 'Desire', 'Admiration', and 'Anger' had low recognition hit rates for intended emotion but higher percentages of emotional experience (above 65%), as shown in Figures 5 and 7. In addition, data analysis identified emotions with critical percentages, sparking a discussion about which emotions are often aroused by visual or physical interactions with products in different cultural contexts. The terms 'Pride', 'Hope', 'Disgust', 'Shame', 'Fear', and 'Contempt' had percentages equal to or lower than 51% for emotional experience with products, being the worst results of 'Shame' (29%) and 'Fear' (33%) even considering the general percentages (Figure 7). Therefore, our results oppose the idea that the 14 emotions included in PrEmo® are essential for evaluating the product experience, as described in previous works (Laurans; Desmet, 2017; Desmet, 2018).

Implications

Based on the results of the intended emotion recognition of the current study, it is suggested that researchers, designers, engineers, or other practitioners take precautions when using PrEmo®, especially in tests conducted in Brazil. During the application of the PrEmo® in the static version (characters printed on cards), in interviews, focus groups, or any research method, the manual states that researchers can ask the participants how they feel about a stimulus, asking them to explain why they feel that way or explaining in their own words what the card expresses, as an optional complementary gathering data. However, based on our results, it is recommended that these questions necessarily be included in the method and its procedures to avoid confusion and noise in the subsequent data interpretation. The card-set version contains no information about the emotion represented by the character, its meaning, or the emotion's name. Therefore, without moderation from the researcher or designer, the participant only gets in touch with emotional cues expressed through facial and body language. Following the procedures suggested above when applying PrEmo® in research can avoid noise in data and interpretation bias, as well as in design solutions.

Furthermore, the low levels of emotional experience raised questions about applying the fourteen PrEmo® characters to evaluate any product or service. An alternative approach to facilitate data interpretation may be to exclude 'Pride', 'Admiration', 'Shame', 'Fear', and 'Contempt'. Another alternative to mitigate bias in data interpretation may be to incorporate new terms associated with emotions to categorize a character, thereby creating layers of interpretation. For example, when analyzing 'Joy' data, researchers should include another layer with the term 'Positive surprise'. This alternative can reduce misunderstandings and expand the analysis of emotional data, considering data collection only with PrEmo® but also triangulation with data collected by different methods (e.g., facial microexpressions). For this enrichment, qualitative methods, such as interviews or focus groups, could be incorporated to complement the quantitative data and provide a deeper understanding of the emotional experiences of participants when interacting with products, which could perhaps contribute to the lack of reliability (or justify it) concerning the intended emotion. For Alves *et al*. (2023), a mixed-methods approach to research allows a more comprehensive understanding of the results.

As Laurans and Desmet (2017) indicated, researchers and other practitioners cannot expect a perfect match between non-verbal representations, such as PrEmo® characters and emotions. According to the authors, achieving recognition rates of 95-100% is impossible when validating any non-verbal self-report instrument. Nonetheless, it is relevant to seek better instrument performance by modifying non-verbal representations, creating new ones, or adapting the instrument's application.

Limitations and Further Research

This study has limitations because of its aims, leaving some opportunities for future research. In line with the Community of Portuguese Language Countries (CPLC), some nations in Europe, America, Africa, and Asia use Portuguese to communicate daily. CPLC is a multilateral forum to deepen cooperation among its nine Member States: Angola, Brazil, Cape Verde, Guinea-Bissau, Equatorial Guinea, Mozambique, Portugal, São Tomé and Príncipe, and East Timor (CPLC,

2021). All Member States are considered Portuguese-speaking because Portuguese is their official or dominant language.

In 2009, the implementation of the New Orthographic Agreement unified the language in CPLC. However, the Portuguese language still presents linguistic variations depending on the culture, such as differences in pronunciation, grammar, and vocabulary. According to the Population Pyramid (2020), Brazil is the Member State with the largest population, around 212,559,409 people (72.7% of the CPLC's population). Therefore, validating PrEmo® in Brazil, with its cultural and idiomatic specificities, was crucial. Therefore, this study was carried out only in Brazil with Brazilian residents. However, it would be interesting to verify whether populations from other countries in the CPLC had similar results. Moreover, other Latin languages widely spoken around the world should advance PrEmo® validation with cross-cultural adaptation, such as Spanish and French. Thus far, the tool has yet to be validated in countries where these languages are spoken.

In this study, we evaluated the PrEmo® card-set version. This version lacks other emotional cues, such as vocal intonation and gestures, which differ from the version with animated characters validated by Laurans and Desmet (2017). Although the tool's manual (Desmet, 2019) indicated a mobile application (optional) to see dynamic animations, it wasn't available at the time of our data collection. Future studies with a Brazilian sample can establish a partnership with Desmet (2019) and verify whether there are significant differences between the static and animated versions of PrEmo®.

The knowledge and previous use of PrEmo® weren't considered an exclusion criterion for the sample of Phase 2. Future research should evaluate previous experience with the tool as a variable. Regarding the sample, the majority had a high educational level, with 61.7% holding a postgraduate degree. It would be possible to expand the current study by including people with lower educational degrees to examine how this influences the emotional recognition rate. Evaluating the influence of the education level on the interpretation of PrEmo® would bring to light relevant information for Methods for emotion measurement in Design, as well as investigating generational and gender differences. Lastly, testing the usefulness of PrEmo® in actual self-report situations would be valuable, as would comparing our findings and verifying whether the same interpretation failures occur.

Conclusion

The current study aimed to conduct a validation process with a cross-cultural adaptation of the non-verbal self-report PrEmo® in Brazil, a country in South America with Portuguese as its official language. It is essential to highlight that this study did not aim to invalidate PrEmo® and the studies that have adopted it as an instrument for collecting emotional data. However, the validation process conducted in this study showed low reliability of the tool.

The results showed that the association between the PrEmo® characters and the intended emotions presented some flaws, with culture as a relevant variable. Although previous studies have validated the tool, the characters representing the emotions 'Pride', 'Joy', 'Fascination', and 'Admiration' still have low recognition rates for the intended emotions. The participants had difficulty identifying the differences between emotional cues in facial and bodily expressions, and therefore, distinguishing these emotions. In summary, to avoid these misunderstandings, it's

recommended that when using PrEmo®, researchers ask participants how they feel about the stimulus, asking them to explain why they feel that way or explaining in their own words what the card expresses. Moreover, it's suggested that designers and other practitioners adopt a mixed-methods approach, combining PrEmo® and qualitative methods to gain a wider view of the phenomenon under investigation and minimize bias in data interpretation. In turn, negative emotions achieved better success rates in emotional recognition than positive emotions, especially 'Disgust', 'Sadness', and 'Contempt', and the participants could attribute the intended valence to the characters.

Based on our findings, we identified which PrEmo® characters require improvement. This insight highlighted the need to redesign some characters and guided the definition of guidelines or strategies to interpret the emotional data collected by the tool more accurately in research with Brazilian individuals. Our findings are relevant for Brazilian designers and other practitioners who perform emotional evaluations of products and services. Future research should validate PrEmo® in other countries of CPLC. Furthermore, this study can contribute to emotion-driven knowledge and to the Methods for measuring emotion in the Design field, as well as designers and human-centered practitioners who work with product development and innovation through an emotion-oriented approach.

Acknowledgments

The authors thank CAPES (Financial Code 001) and the participants for their time.

References

ALVES, A. L.; MARTELI, L. N.; SILVA, F. M.; PASCHOARELLI, L. C. Design, Emotion, and Color: Analysis Study of the Scientific Production in the International Color Association Proceedings. In: BROEGA, A. C.; CUNHA, J.; CARVALHO, H.; PROVIDÊNCIA, B. **Advances in Fashion and Design Research**. CIMODE 2022. Cham: Springer, 2023. Cap. 7, p. 759-770.

BÄNZIGER, T.; TRAN, V.; SCHERER, K. R. The Geneva Emotion Wheel: A tool for the verbal report of emotional reactions. In: **Proceedings of Conference of the International Society of Research on Emotion**, 2005, p. 271-294.

BEATON, D. E.; BOMBARDIER, C.; GUILLEMIN, F.; FERRAZ, M. B. Guidelines for the process of cross-cultural adaptation of self-report measures. **Spine**, v. 25, n. 24, p. 3186-3191, 15 December 2000. Monthly.

BRASIL. Censo Demográfico. **Instituto Brasileiro de Geografia e Estatística**, 2022. Available at: https://www.ibge.gov.br/en/statistics/social/population/18391-2010-populationcensus.html?=&t=resultados. Accessed on 3 September 2022.

BROEKENS, J.; BRINKMAN, W. AffectButton: A method for reliable and valid affective self-report. **International Journal of Human-Computer Studies**, v. 71, n. 6, p. 641-667, 01 June 2013. Monthly.

CARDELLO, A. V.; JAEGER, S. R. Measurement of consumer product emotions using questionnaires. In: MEISELMAN, H. L. **Emotion measurement**. Amsterdam: Woodhead Publishing, 2021. Cap. 2, p. 273-321.

CPLP. Estados Membros. **Comunidade dos Países de Língua Portuguesa**, 2023. Available at: <https://www.cplp.org/id-2597.aspx>. Accessed on: 10 June 2023.

DAMÁSIO, A. R. **The Strange Order of Things:** Life, Feeling and the Making of Cultures. New York: Pantheon, 2021. 336 pages.

DAMAZIO, V.; TONETTO, L. M. Design Emocional e Design para o Bem referências e apontamentos Design Emocional e Design para o Bem-Estar: marcos, referências e apontamentos. **Estudos em Design**, Rio de Janeiro, v. 30, n. 1, p. 156-170, 2022. Four-monthly.

DARWIN, C. **The Expression of the Emotions in Man and Animals**. Ed. Francis Darwin, 2ª ed. rev. Cambridge: Cambridge University Press, 2009. 394 pages.

DESMET, P. M. A. Measuring emotion: Development and application of an instrument to measure emotional responses to products. In: BLYTHE, M.; MONK, A. **Funology 2:** From Usability to Enjoyment. Cham: Springer, 2018. Cap. 26, p. 391-404.

DESMET, P. M. A. **PrEmo (Emotion Measurement Instrument) card set:** Male version. Amsterdam: Delft University of Technology, 2019. 2 pages.

DUARTE, S.; CUNHA, J.; BUHAMRA, C.; QUENTAL, J. Design e emoção: a utilização da ferramenta PrEMO para testes em projeto de vitrina popular. In: **Anais** do 14º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design. São Paulo: Blucher, 2022. p. 163-179.

FOKKINGA, S. F.; DESMET, P. M. A. Emotion Typology. **Delft University of Technology**, 2022. Available at: https://emotiontypology.com. Accessed on: 10 January 2023.

GOBBI, A. G.; CATECATI, T.; MERINO, E. A. D.; MERINO, G. S. A. D.; FERREIRA, M. G. G. Uso do eye tracking para medição da satisfação para testes de usabilidade em interfaces web. In: **Anais** do 16° USIHC – Congresso Internacional de Ergonomia e Usabilidade de Interfaces Humano Computador. São Paulo: Blucher, 2017. p. 2287-2296.

HESS, U.; THIBAULT. P. Darwin and emotion expression**. American Psychologist**, Washington, v. 64, n. 2. p. 120-128, February-March 2009. Nine times a year.

HOFSTEDE, G.; HOFSTEDE, G.; MINKOV, M. **Cultures and organizations:** software of the mind: intercultural cooperation and its importance for survival. London: McGraw-Hill, 2010. 561 pages.

JAEGER, S. R.; ARES, G. Dominant meanings of facial emoji: Insights from Chinese consumers and comparison with meanings from internet resources. **Food Quality and Preference**, Amsterdam, v. 62, p. 275-283, December 2017. Monthly.

LAURANS, G.; DESMET, P. M. A. Developing 14 animated characters for non-verbal self-report of categorical emotions. **Journal of Design Research**, Geneva, v. 15, n. 3-4, p. 214-233, 19 September 2017. Four-monthly.

MOKKINK, L. B.; PRINSEN, C. A.; PATRICK, D. L.; ALONSO, J.; BOUTER, L. M.; DE VET, H. C.; TERWEE, C. B. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. Amsterdam: University Medical Centers, 2019. 132 pages.

OZOLINS, U.; HALE, S.; CHENG, X.; HYATT, A.; SCHOFIELD, P. Translation and back-translation methodology in health research – a critique. **Expert review of Pharmacoeconomics & Outcomes research**, Abingdon, v. 20, n. 1, p. 69-77, 28 February 2020. Ten times a year.

PASCHOARELLI, L. C.; SILVA, F. M.; SILVA, J. M. N. Colour in product design and its influence on User Experience (UX) and Usability (US): a cross-sectional study. **Estudos em Design**, Rio de Janeiro, v. 32, n. 3, p. 180-192, 2024. Four-monthly.

PENG, Z.; DESMET, P. M. A.; XUE, H. Mood in Experience Design: A Scoping Review. She Ji: **The Journal of Design, Economics, and Innovation**, Shanghai, v. 9, n. 3, p. 330-378, 04 December 2023. Four-monthly.

PLUTCHIK, R. **Emotions and Life:** Perspectives from Psychology, Biology, and Evolution. Washington: American Psychological Association, 2002. 592 pages.

POPULATION PYRAMID. List of countries ordered by population size. **Population Pyramid**, 2020. Available at: https://www.populationpyramid.net/pt/popula%C3%A7%C3%A3o/2020/. Accessed on: 24 January 2024.

PORGES, S. W. **The polyvagal theory:** Neurophysiological foundations of emotions, attachment, communication, and self-regulation. New York: WW Norton & Company, 2011. 272 pages.

SANTOS, L.; ALENCAR, N.; RODRIGUES, J.; GONÇALVES, A.; TROMBETE, F. Proposal of two facial scales using emojis as tools to understand emotions in research with Brazilian consumers. **Acta Scientiarum:** Technology, Maringá, v. 44, p. e56610, 12 January 2022. Annual.

SCHERER, K. R. What are emotions? And how can they be measured? **Social science information**, Paris, v. 44, n. 4, p. 695-729, December 2005. Quarterly.

YOON, J. K.; POHLMEYER, A. E.; DESMET, P. M.; KIM, C. Designing for positive emotions: issues and emerging research directions. **The Design Journal**, London, v. 24, n. 2, p. 167-187, 19 November 2020. Bimonthly.

YOUNGSTROM, E. A.; GREEN; K. W. Reliability generalization of self-report of emotions when using the differential emotions scale. **Educational and Psychological Measurement**, Englewood Cliffs, v. 63, n. 2, p. 279-295, 01 April 2003. Bimonthly.

About the Authors

Ana Laura Alves
Researcher at the University of São Paulo (USP). PhD in Design at the São Paulo State University (UNESP). Specialist in Ergonomics at the State University of Londrina (UEL) and certified ergonomist by the Brazilian Ergonomics Association (ABERGO). Master's and bachelor's degrees in Design from São Paulo State University. Member of ProCor do Brasil Association and ABERGO. ORCID. https://orcid.org/0000-0002-8367-5796

Gerusa Ferreira Lourenço
Associate Professor at the Department of Occupational Therapy at Federal University of São Carlos (UFSCar). Post-doctorate in Design, São Paulo State University. PhD and Master's in Special Education, PPGEEs/UFSCar. Degree in Occupational Therapy, UFSCar. Professor of the Occupational Therapy course and the Postgraduate Program (master's and doctorate's) in Special Education at UFSCar. ORCID. https://orcid.org/0000-0002-7550-4858

Luis Carlos Paschoarelli
Full Professor at the Department of Design (UNESP). Post-doctorate in Inclusive Design (ULisboa) and in Ergonomics (UTLisboa); PhD in Production Engineering (UFSCar). He is a leader at the 'Ergonomic Design: Project and Interfaces - Research Group' (dgp.cnpq.br/dgp/espelhogrupo/18099) and coordinator of the Laboratory of Ergonomics and Interface (pnipe.mcti.gov.br/laboratory/12334). ORCID. https://orcid.org/0000-0002-4685-0508