doi: https://doi.org/10.15446/rcp.v32n2.99016

Validation of the Hispanic American Version of the Plymouth Sensory Imagery Questionnaire (psi-q): A Culturally Adapted Measure of Multisensory Mental Imagery

Running Head: Validation of the Hispanic American PSI-Q

IONA NAISMITH
Department of Psychology, Universidad de los Andes, Bogotá, Colombia

WILLIAM JIMÉNEZ LEAL
Department of Psychology, Universidad de los Andes, Bogotá, Colombia

ASHLEY POPLE
Department of Economics, University of Oxford, United Kingdom

EMILY A. HOLMES
Department of Psychology, Uppsala University, Uppsala, Sweden
Department of Clinical Neuroscience, Karolinska Institutet, Sweden

How to cite this article: Naismith, I., Jiménez-Leal, W., Pople, A., & Holmes, E., (2023). Validation of the Hispanic American version of the Plymouth Sensory Imagery. *Revista Colombiana de Psicología*, *32*(2), 51-63. https://doi.org/10.15446/rcp.v32n2.99016

Correspondence: Iona Naismith (https://orcid.org/0000-0003-3897-783X); Universidad de los Andes, Carrera 1 # 18A-12, Bogotá, Colombia. Telephone: 337499, Ext. 2547. E-mail: ij.naismith@uniandes.edu.co

SCIENTIFIC RESEARCH ARTICLE

RECEIVED: OCTOBER 13TH, 2021 - ACCEPTED: JANUARY 1ST, 2023

Abstract

Mental imagery is increasingly recognized to play a key role in psychotherapy, education, and other domains. The Plymouth Sensory Imagery Questionnaire (PSI-Q) measures the vividness of mental imagery in seven modalities: vision, sound, smell, taste, touch, bodily sensations, and feelings. This study adapted the PSI-Q for Hispanic-American respondents in Colombia and explored moderators of imagery vividness. Study 1 validated the original PSI-Q (short version) in a sample of 292 Colombian university students. Study 2 developed and validated a cultural adaptation of the PSI-Q with 508 Colombian adults from the general population. The Vividness of Visual Imagery Questionnaire-2 (VVIQ-2) was used to analyze convergent validity in both samples. Thus we propose a new 29-item Hispanic-American version of the PSI-Q, adapted to cultural differences in the region, that displays good reliability and convergent validity. Imagery vividness was higher in females and at higher education levels. Higher cross-loadings between senses occurred in this sample, which may reflect cultural differences in somatization.

Keywords: Assessment, Hispanic, Latin America, mental imagery.

Validación de la Versión Hispanoamericana del Cuestionario de Imaginería Sensorial de Plymouth (psi-q): Una Medida Culturalmente Adaptada de la Imaginería Mental Multisensorial

Resumen

Se reconoce cada vez más que las imágenes mentales tienen un papel clave en la psicoterapia, la educación y otros dominios. El Cuestionario de Imágenes Sensoriales de Plymouth (PSI-Q) mide la viveza de la imaginería mental en siete modalidades sensoriales: visión, sonido, olfato, gusto, tacto, sensaciones corporales y sentimientos. El presente estudio adaptó el PSI-Q para hispanoamericanos en Colombia y exploró moderadores de percepción de las imágenes. El estudio 1 validó el PSI-Q original (la versión corta) en una muestra de 292 estudiantes universitarios colombianos. El estudio 2 desarrolló y validó una adaptación cultural del PSI-Q con 508 adultos colombianos de la población general. Se utilizó el Cuestionario de Vividness of Visual Imagery-2 (VVIQ-2) para analizar la validez convergente en ambas muestras. Proponemos una nueva versión hispanoamericana de 29 ítems del PSI-Q, que se adapta a las diferencias culturales de esta región y muestra buena confiabilidad y validez convergente. La viveza de las imágenes fue mayor en las mujeres y en los niveles educativos más altos. Se observaron cargas cruzadas altas entre los sentidos, lo que puede reflejar diferencias culturales en la somatización.

Palabras clave: América Latina, evaluación, hispánico, imágenes mentales.

Mental imagery subjectively resembles the experience of perceiving a stimulus, but it occurs in the absence of the corresponding external stimulus (Reisberg et al., 2003; Pearson et al., 2015). Although the Visual modality has been the focus of most empirical studies on imagery, mental imagery can occur in any of the sensory modalities. Images can be involuntary or deliberately generated and may involve memories or imagined hypothetical scenarios.

Mental images can be measured according to their vividness: how closely they resemble real perceptual experience. Vividness is influenced by the availability of cognitive processes, e.g., concurrent cognitive load in the so-called visuospatial sketchpad of working memory (Baddeley & Andrade, 2000). It is also shaped by individual differences (e.g., in executive functions and possibly the perceptual cortices) and by the qualities of the stimuli being imagined, including their complexity (Bywaters et al., 2004). The importance of measuring imagery vividness is supported by multiple studies that have shown that vivid mental images powerfully impact our emotions and, when negative, can be associated with psychopathology (Holmes et al., 2008; Holmes & Mathews, 2010). Vivid imagery is associated with desire and motivation (Kavanagh et al., 2005), motor performance (Callow et al., 2006), and problem-solving success (Kozhevnikov et al., 2007). However, most measures of mental imagery vividness only assess the Visual modality (Marks, 1995).

Andrade et al. (2014) developed the Plymouth Sensory Imagery Questionnaire (PSI-Q), a self-report scale that measures the vividness of mental imagery in seven modalities: Vision, Sound, Smell, Taste, Touch, Bodily Sensation, and Emotional Feeling. The original paper presents an extended 35-item PSI-Q (5 items per subscale), as well as a shorter 21-item version (3 items per subscale). The PSI-Q differs from most previous imagery scales by measuring multiple sensory modalities simultaneously. It enables a comparison of relative strengths and weaknesses across modalities at a group and individual level and helps to identify the modality-general and modality-specific contributions to imagery (McNorgan, 2012). Studies have used the PSI-Q to demonstrate a relationship between sensory imagery and sensory sensitivity (Dance et al., 2021) and how imagery in different modalities changes through development (Arshamian et al., 2020).

The PSI-Q has recently been validated in Spain (Pérez-Fabello & Campos, 2020) and Germany (Jungmann et al., 2022). Both studies found support for the 7-factor model and good psychometric properties using the original set of items. However, to date it has only been validated in high-income countries, despite the fact that only 16% of the world's population lives in such regions (Prydz & Wadhwa, 2019). Consequently, this has an impact on the breadth of research that can be done on mental imagery and its potential role in shaping human motivation, emotion, and learning. Therefore, the present study aims to adapt and validate the PSI-Q for Hispanic America, the Spanish-speaking nations of the Americas. These countries share a common language and many cultural similarities, so instruments can often be used across the region. Comprising predominantly of middle-income countries, the region contains 433 million people, representing 5.7% of the world population (World Bank, 2019). Thus, this is an important region in which to support research and informed clinical practice that is more representative of human functioning globally.

In study 1, the psychometric properties of a simple (non-adapted) translation of the PSI-Q short version were explored in a sample of 292 private university students in Colombia. The short version was chosen for its excellent psychometric properties in English and because of its wide use by researchers due to convenience. Convergent validity was analyzed using the Vividness of Visual Imagery Questionnaire-2 (VVIQ-2), the most widely used instrument of mental imagery vividness, and test-retest reliability was also calculated. In study 2, the long version of the PSI-Q was culturally adapted for Hispanic America (by adjusting item wording, instructions, and response scale anchors, and replacing some items), and then applied in a

community adult sample of 508. Convergent validity using the VVIQ was again calculated.

Study 1

Materials and Methods

Participants

Participants were 292 Colombian undergraduate psychology students from a private university in Bogotá (210 female, 71.9%). Age ranged from 18 to 34 years (M = 20.3). Sample size is consistent with guidelines recommending an approximate sample size of *N* = 300 for a factor analysis (Field, 2005; Tabachnick & Fidell, 2012).

Measures

Plymouth Sensory Imagery Questionnaire (PSI-Q). For this study, the short version of the test was used following direct translation into Spanish (see Procedure).

Vividness of Visual Imagery Questionnaire-2 (VVIQ-2; Marks, 1995). The VVIQ-2 consists of four sets of four items. Each set asks respondents to imagine a particular visual scene with their eyes open and then answer four questions about the vividness of elements of their image on a scale from 1 (no image at all, you only "know" that you are thinking of an object) to 5 (perfectly clear and as vivid as normal vision). For example, one set of items asks participants to "visualize the rising sun" and then report the vividness of four elements, including "the sun is rising above the horizon into a hazy sky" and "a rainbow appears". The present study used the validated Spanish version (Soledad Beato et al., 2006).

Procedure

Ethical consent was obtained from the ethics committee in the university at which this study was conducted. Following established guidelines (Sousa & Rojjanasrirat, 2010), a translation of the PSI-Q from English to Spanish was prepared by two bilingual and bicultural individuals whose mother tongue is Spanish (the second author and a postgraduate student). The two versions were compared by another student and discrepancies were resolved as a team. Back translation was then carried out by the first author, whose first language is English, and ambiguities were revised and resolved by the team. At this stage, items were not adapted or added because (i) we wanted to explore the properties of the original PSI-Q, and (ii) all items were deemed familiar and comprehensible to a private-university student population, who have been exposed to international cuisine, activities, etc. Recruitment took place in psychology lectures. All data was collected via pen-and-paper. In the initial session, all participants completed a consent form and the PSI-Q (short version), and a subgroup of N = 160 completed the VVIQ for convergent validity. Participants were given a snack to thank them for participating.

Six weeks later, the PSI-Q (short version) was reapplied to 65 participants (two classes, selected randomly) to measure test-retest reliability. Age (M = 19.94, SD = 2.86) and gender (49 females, 75.4%) were similar to that of the original sample. The data is available through open-access repository [https://osf.io/qrb6j/].

Data analysis

Analysis was conducted using RStudio version 1.1456 (RStudio Team, 2016) with the Lavaan library (Rosseel, 2012). First, we ran a confirmatory factor analysis and checked goodness-of-fit indices and factor loadings to examine whether the factor structure replicated that of the English version. We then proceeded to adjust this model to improve the overall fit. Notice that for the confirmatory factor analysis we have used one of the robust maximum likelihood estimators (MLF) which allowed us to run the analysis with incomplete data.

Results

Items showed mostly positive asymmetric distributions, with neither floor nor ceiling effects. Means and standard deviations (in brackets) for

each subscale are as follows: Visual = 7.92 (1.51), Auditory = 8.01 (1.91), Olfactory = 5.51 (2.53), Taste = 6.57 (2.34), Tactile = 7.51 (1.99), Bodily Sensations = 7.16 (2.02), and Emotion = 7.80 (1.85).

Internal consistency

Internal consistency was excellent for the PSI-Q total (α = .90, ω = .91) and acceptable to good for Visual (α = .78, ω = .79), Auditory (α =. 60, ω = .62), Olfactory (α = .75, ω = .77), Taste, (α = .73, ω = .75), Tactile (α = .71, ω = .72), Bodily Sensations (α=.68, ω = .68), and Emotion scales (α = .70, ω = .71). Mean correlation between the scales was of 0.53, ranging from 0.32 ($p < 0.01$) for the correlation between Taste and Vision to 0.6 ($p < 0.01$), for the correlation between Taste and Smell.

Convergent validity with the VVIQ

The VVIQ scale exhibited a good reliability (α =. 85, ω = .88). Moderate convergent validity was seen between the VVIQ and the original PSI-Q full scale ($r = 0.53$). All PSI-Q subscale scores were positively associated with the VVIQ, with the Visual and Bodily Sensations subscales having the highest and lowest correlations ($r = .67$ and $r = .31$), respectively.

Test re-test reliability

Eight-week test-retest reliability, measured with 65 participants, indicated a significant correlation ($r = .65$), with no significant differences between the means of both applications (M*pre* = 256.78, SD = 26.3, M*post* = 251.98, SD = 26.3), paired *t*-test: $t(64) = 0.73$, $p = .46$, M*diff* = 4.8, 95% CI [-8.4, 18.01]).

Confirmatory factor analysis of short version

A confirmatory factor analysis offered adequate support for the original model of seven distinct scales. Although the model fit is acceptable for a replication, it was further improved by removing two items with potentially high cross-loadings, "burnt wood" (Smell) and "walking briskly in the cold" (Bodily Sensation), and allowing for residual covariances that include the Smell scale items (See Table 1 for a summary). This may reflect that experiences of wood-burning and cold weather are uncommon in Colombia. The residual covariances with Smell scale items stemmed from high cross-loadings between Smell and other subscales, particularly Taste, resembling findings of the original PSI-Q validation study (Andrade et al., 2014).

The best-fitting model performs better in every single indicator assessed, with better standard fit indices, like the CFI (0.94) and TLI (0.92), and a good RMSEA (0.050). Full factor loadings for the best model are presented in Electronic Supplementary Material 1.

Table 1

Contrast of Goodness-of-fit Measures for Structural Equation Modelling (SEM) Models for the PSI-Q (N =292)

	Original model	Best model
Number of parameters	63	59
χ2	384.34 (168), p < 0.001	218.52 (127), p < 0.001
CFI	0.896	0.9475
TLI	0.870	0.928
Log likelihood	-13115.55	-11807.89
AIC	26357.09	23704.29
BIC	26587.64	23953.84
RMSEA	0.067	0.050
CI Lower	0.058	0.037
CI Upper	0.076	0.061
P value	0.001	0.090
GFI	0.882	0.923

Note. The GFI refers to the goodness-of-fit index (Sharma et al., 2005) and it is analogous to *R2* in the context of linear regression.

Study 2

Material and Methods

Participants

Ethical consent for the second study was obtained from the institutional IRB at Universidad de los Andes (ref #14638). Participants in the second study comprise low-income adults based in Bogotá who had signed up for an entrepreneurship training program that incorporated aspects of mental imagery. All participants completed a consent form. The data is available through open-access repository [https://osf.io/qrb6j/].

The present study used data collected at baseline on two occasions. First, in April 2019, 43 participants (22 male, 21 female) between 19 and 62 years old (M = 40.56, SD = 13.68) were recruited for a pilot phase of the entrepreneurship training program. Second, in June 2019, 508 adults (261 women, 259 men) were recruited for the main phase of the same project. These participants represent a relatively varied sample of the urban Colombian population. The majority (62.3%) were aged 18-28 years, with 31.8% aged 29-59 and 5.9% aged 60+. In terms of education level, 187 (36.8%) had at least an undergraduate university degree, 172 (33.8%) had a technical (non-university) qualification, 91 (18.1%) had completed high-school, and 58 (11.4%) had not completed high-school. It is worth mentioning that the entrepreneurship program prioritized recruitment of vulnerable individuals, such as victims of the Colombian armed conflict, migrants, and low-income households. However, PSI-Q subscale means in the present study were very similar to those found in the original PSI-Q validation (Andrade et al., 2014), suggesting that participants did not suppress imagery when completing the PSI-Q.

Measures

Plymouth Sensory Imagery Questionnaire (PSI-Q).

A direct translation of the original 35-item instrument was applied in the pilot phase (as in Study 1), but some participants identified that several items were difficult to understand. Consequently, the scale was adapted for context, with four of the items removed and replaced with five new items, and one further item reworded (see Procedure). The full translated version is provided at the end of this article.

Vividness of Visual Imagery Questionnaire-2 (VVIQ-2; Marks, 1995).

See Study 1 for details.

Procedure

The team who adapted the PSI-Q consisted of bilingual Colombian and British researchers living in Colombia, experienced in implementing interventions with participants from across the spectrum of socio-economic status.

In the pilot, participants completed the validated Spanish version of the PSI-Q, as in Study 1. Items were administered verbally via one-on-one interviews by trained research assistants to enable evaluation of participants' comprehension item-by-item. Research assistants read instructions and items aloud, and visually presented the response scale each time (see final version at the end of this article). The research assistants identified any items that had been difficult for participants to understand.

A new version of the PSI-Q adapted for a Hispanic American context was then developed based upon the insights of the pilot (see results) and applied in the second phase. Data collection used the same approach as the pilot (one-on-one verbal administration) to ensure that low literacy did not affect validity. Researchers supervising data collection ran programmed quality checks daily to correct for duplicates and identify any deviations from the protocols. They were also in regular contact with field staff to clarify possible misunderstandings and conducted random spot checks to ensure that surveys were being conducted according to protocols.

Data analysis

As in Study 1, analysis was conducted using RStudio version 1.1456 (RStudio Team, 2016) and confirmatory factor analyses ran with the Lavaan library (Rosseel, 2012) with an ML estimator. We fitted several models to contrast the original, shortened, and modified versions of the PSI-Q, and explored some moderators of imagery vividness.

Results

Results from pilot study and subsequent adaptation

Mean vividness scores indicated satisfactory ability to imagine (≥ 7 of 10) for most items, except for a "stuffy room" (Smell) with mean vividness of 6.6. As in Study 1, items showed a positive asymmetrical distribution that suggests the task was relatively easy for the participants.

Through group discussions and analysis of pilot data, several of the PSI-Q items were hypothesized to have low face validity for a representative sample of Hispanic American population. Specifically, these were: "mustard" and "black pepper" (Taste), since these condiments are not commonly used in Hispanic American cuisine; "relaxing in a warm bath" (Bodily Sensation), since baths are rarely found outside upper-class homes; and "walking briskly in the cold" (Bodily Sensation), since most countries in Hispanic America have a warm or temperate climate year round. These four items were therefore removed and replaced with more contextually relevant items.

The item "newly cut grass" (Smell) was reworded to "damp grass". In Hispanic America, grass tends to be found in communal parks or areas, but not in private gardens, and thus people rarely cut grass.

An alternative item pool for the subscales Taste, Smell and Bodily Sensations was developed and refined through group discussions among the research team. We retained only those items deemed most comprehensible and familiar to participants across Hispanic America from all socio-economic backgrounds. The new items included for testing were: "car exhaust fumes" (Smell); "mint leaves" (Taste); "orange juice" (Taste); "sitting on a hard wooden chair" (Bodily Sensation); and "your leg falling asleep" (Bodily Sensation).

The scale instructions were lengthened in detail to ensure that respondents understood what was being asked of them. Additionally, the response scale was edited to assist comprehension by (i) shading the scale in darkening shades of grey from left to right, and (ii) using additional word anchors to clarify the meaning of the scale values (see the final version at the end of this article for instructions and response scale).

Descriptive statistics of adapted PSI-Q

Table 2 presents means and standard deviations for the individual items, scale and subscale totals for the full sample of N = 508 who completed the adapted 36-item scale. The items with comparatively low vividness (M < 7) were: "a stuffy room" (Smell), "burnt wood" (Smell), "warm sand" (Touch), and "scared" (Feeling).

Table 2
Descriptive Statistics for the Adapted Version of the PSI-Q (N=508)

Dimension	M	SD
Imagine the appearance of:	8.34	2.18
A friend you know well	8.3	2.2
A cat climbing a tree	8.2	2.3
A sunset	8.7	1.9
The door in front of your house	9.0	1.8
A fire	7.5	2.7

Dimension	M	SD
Imagine the sound of:	8.4	2.18
An ambulance siren	8.0	2.4
Hands clapping in applause	8.5	2.0
The mewing of a cat	8.4	2.3
The sound of a car horn	8.5	2.2
The sound of children playing	8.6	2.0
Imagine the smell of:	7.32	2.8
A stuffy room	6.8	3.1
A rose	7.4	2.8
Fresh paint	8.2	2.4
Damp grass	7.4	2.7
Car exhaust fumes	7.3	2.9
Burnt wood	6.8	2.9
Imagine the taste of:	8.12	2.28
Mint	7.2	3
Toothpaste	8.7	1.8
Lemon	8.7	1.9
Salty water	7.4	2.8
Orange juice	8.6	1.9
Imagine touching:	7.76	2.52
Warm sand	6.4	3.1
A soft towel	7.9	2.3
The point of a pin	7.5	2.9
Icy water	8.5	2.0
Fur	8.5	2.3
Imagine the bodily sensation of:	7.76	2.54
Sitting down on a hard wooden chair	7.8	2.5
Sore throat	7.6	2.6
Threading a needle	7.5	2.7
Jumping in a pool	7.7	2.6
Having your leg fall asleep	8.2	2.3
Imagine feeling:	8.04	2.3
Excited	9.0	1.6
Relieved	8.8	1.8
Furious	7.5	2.7
In love	8.1	2.5
Scared	6.8	2.8

Values for Cohen's alpha and McDonald's omega, classic measures of internal reliability, reveal acceptable to good internal consistency, assuming unidimensionality, for all subscales of the PSI-Q (see Table 3). Consistency did not improve with the removal of any item, with the exception of "jumping into a pool", which marginally improved reliability of Bodily Sensations to $\alpha = 0.75$, $\omega = 0.76$. Correlations between the subscales are moderate except for the Emotion subscale, which has the lowest correlations with the other dimensions (See Table 3).

Table 3
Reliability Measures for Subscales of the PSI-Q for Study 2 and Correlations Between Subscales (N=508)

PSI-QSubscale	mean	SD	alpha	omega	Visual	Auditory	Olfactory	Taste	Tactile	Bodily S.
Visual	8.34	2.18	0.82	0.83	--	--	--	--	--	--
Auditory	8.40	2.18	0.74	0.75	0.62	--	--	--	--	--
Olfactory	7.32	2.80	0.75	0.76	0.52	0.60	--	--	--	--
Taste	8.12	2.28	0.69	0.72	0.44	0.53	0.66	--	--	--
Tactile	7.76	2.52	0.76	0.79	0.49	0.56	0.63	0.64	--	--
Bodily sensations	7.76	2.54	0.72	0.73	0.43	0.54	0.60	0.54	0.64	--
Emotion	8.04	2.30	0.83	0.84	0.40	0.47	0.43	0.47	0.48	0.54

Note: All correlations are significant at $p < 0.01$.

The VVIQ-2 scale exhibited a good reliability (α =. 87, ω = 0.90). As before, moderate convergent validity was seen between the VVIQ-2 and the PSI-Q full scale (r=0.57). All PSI-Q subscale scores were positively associated with the VVIQ-2, with the Taste and Feeling subscales having the highest (r = .52) and lowest correlations (r = .39) respectively.

Confirmatory factor analysis

An initial confirmatory factor analysis showed an acceptable fit for replication, with an RMSEA of 0.07 and a SRMR of 0.067 (See Table 3). The model fit, however, improves significantly with various modifications, as outlined below.

Modification indices suggested fitting a model including residual covariances between items in the Feelings scale and between the items "car fumes" and "burnt wood". This model, also presented in Table 3 (third column), is the best fitting model without allowing item cross-loading. However, this model still does not exhibit a good fit, with CFI and TLI indices under .90 and an RMSEA confidence interval that does not include 0.05.

The modification indices of the model fitted to the original structure suggest several cross-loadings, particularly of Feeling items ("furious", "excited", "scared") onto the Touch and Bodily Sensations scales. While allowing for cross-loadings improves the fit of the simple structure imposed by the analysis, it limits the interpretation of the relevant scores as measures of the constructs at hand. Alternatively, shortening the scale can help to reduce the error variance (Eriksson & Boman, 2018) while maintaining the integrity of the subscales' scores. We fitted a model with a reduced number of items, by deleting those with factor loadings < .35; specifically, from the subscales Touch (items "fur" and "soft towel"), Bodily Sensations ("jumping in a pool"), Smell ("rose"), Taste ("salty water") and Feelings ("furious" and "scared").

The indices for both the model with cross-loadings and the shortened 29-item version are presented in Table 3 (fourth and fifth column). Both models' indices show an improvement relative to the original model and to the best-fitting model without cross-loadings, with the shortened model having the best fit (see Electronic Supplementary Material 2 for factor loadings for this model).

As a result, the version of the PSI-Q that best captures the underlying imaginal dimensions in a Hispanic American context is the Shortened Version, which exhibits the best overall fit (See Table 4), acceptable loading for all items included (See Electronic Supplementary material 2), and includes at least 3 items per dimension (included at the end

of this article). The main difference with the original version lies in the measurement of the Feelings and Touch dimensions, whose items did not seem to uniquely capture the dimension intended.

Table 4
Contrast of Goodness-of-fit Measures for SEM Models for the PSI-Q (N = 508)

	Original model	Best model without cross-loadings	Best model with cross-loadings	Shortened version model
Number of parameters	94	96	106	81
χ2	1833.71(572), p < 0.001	1517.56(534), p < 0.001	7278.68(630), p < 0.01	836.03(354), p < 0.01
CFI	0.810	0.849	0.861	0.910
TLI	0.791	0.832	0.844	0.895
Log likelihood	-34577.14	-33566.20	-34401.1	-28078.09
AIC	69342.82	67324.49	69014.21	56318.20
BIC	69729.30	67719.62	69450.03	56653.35
RMSEA	0.070	0.063	0.603	0.054
CI Lower	0.066	0.060	0.056	0.049
CI Upper	0.074	0.067	0.064	0.059
P value	< .001	< .001	< .001	0.07
GFI	0.79	0.822	0.829	0.890
SRMR	0.067	0.071	0.055	0.049

Moderators of imagery vividness

To determine if there were significant moderators of imagery vividness, we fitted a linear model with age, gender (after eliminating non-responders), and educational level as predictors of the scores in all items. The model ($R2$ = 0.014. F(7, 3527) = 8.35, p < 0.0001) revealed differences between men and women and an interaction between gender and education. Women scored slightly higher than men in overall scores, particularly in the Feelings (MMen = 7.90, MWomen = 8.16, Welch t (496) = -1.83, p = 0.06) and Taste subscales (MMen = 7.86, MWomen = 8.40, Welch t (483) = -3.66, p < 0.001). The interaction between gender and education revealed that the link between visualization and education is stronger for men than for women (t (3541) = -2.13, p = 0.021); in men (but not women) higher formal education is associated with stronger imagery vividness. No other terms or predictors were significant.

General Discussion

The objective of this study was to investigate the psychometric properties of a translated version of the PSI-Q in Colombia and design a final version that could be used across a more varied sample in Hispanic America.

Study 1 indicated that in a highly educated student population, most items of the short version may be appropriate, with the exception of the items "burnt wood" (Smell) and "walking briskly in the cold" (Bodily Sensation). Subscale means and internal consistencies were largely comparable to the original English version. Test-retest reliability was similar to that found for the original English version (.65 versus .71). Thus, when working with highly educated populations such as university students, it may be appropriate to make only minor adaptations to the original PSI-Q.

Study 2 involved a more socioeconomically varied and thus more representative sample of Hispanic Americans recruited from vulnerable groups in particular. Analysis suggested that in order to obtain valid data, a cultural adaptation of the PSI-Q would be appropriate (involving adding or replacing items, rather than a simple direct translation).

Following a trial validation with a 36-item scale, it was concluded that seven items should be dropped to ensure good fit of a 7-factor model. We propose that future imagery research in Hispanic America use the new 29-item Hispanic American PSI-Q. This contains ≥ 3 items per subscale and has good internal consistency.

The removal of 7 items was largely due to cross-loadings of Feeling items, particularly "furious" and "scared", onto the Touch and Bodily Sensations subscales from the full scale. These cross-loadings may reflect cultural differences in subjective experience of emotions: previous literature has identified greater somatization of emotional distress in Hispanic Americans compared to non-Hispanic Americans and Europeans (Dunlop et al., 2019; Gureje et al., 1997; Tófoli et al., 2011), which may reflect cultural and linguistic norms for emotional expression. Thus, we hypothesize that Latin American individuals with non-vivid Touch and Bodily Sensation imagery may also rate their imagery of emotions as non-vivid. Future analysis using item differential response techniques on samples from different cultures could enable further exploration of this hypothesis, in conjunction with qualitative discussions with people who have lived experience to reflect on this.

Imagery vividness was moderated by gender, with women reporting more vivid imagery for Tastes and Feelings. The latter may reflect socialization of gender-related display rules for emotion expression, as many cultures socialize men to inhibit most emotions (Chaplin & Aldao, 2013), and Colombia is no exception (Velandia-Morales & Rincón, 2013). Furthermore, vividness was reported as higher in more highly educated men compared to lower-educated men. No result was seen in women, which might reflect gender differences in careers and technical specialization.

Limitations

Although the scale was adapted with a consideration for the language and culture of all Hispanic American countries, it was only validated in Colombian respondents, albeit across wide ranging education and income levels. Thus, future studies should validate it in other countries. The present study used the VVIQ for covalidation, following the original PSI-Q validation, but future studies might also validate with other measures such as the Spontaneous Use of Imagery Scale (SUIS) (Reisberg, Pearson, & Kosslyn, 2003). The item "car exhaust" performed effectively but may be less relevant in future as fossil fuels are eliminated.

Study 2 involved verbal administration of items due to limited literacy in some participants. Whilst this enabled us to validate the questionnaire in a population often excluded from research and we attempted to mitigate participant demand effects through our instructions, it is nonetheless possible that this could create demand effects. On the other hand, some researchers have found increased consistency in responses to verbally delivered questions, perhaps because this leads to greater reflection prior to responding (Perry et al., 2002). Verbal administration (compared to written) might place different cognitive demands on participants, although experimental studies comparing this are lacking (Bowling, 2005).

Conclusions

Mental imagery has assumed increasing importance in psychological therapies, educational programs, entrepreneurship and other domains. The present study proposes a cultural adaptation of the PSI-Q for Hispanic American regions. Reliability and validity were found to be satisfactory in a representative sample of Colombian adults. It is hoped that the adapted 29-item scale will encourage multisensory imagery research across a broader range of geographical regions, cultures, and education levels.

Hispanic American version of the Plymouth Sensory Imagery Questionnaire

Note: This version of the scale includes all 36 items applied to the full sample of Study 2. The Shortened Version, which we recommend based on our findings, involves dropping the seven asterisked items.

Cuestionario 'Plymouth' de Visualización Sensorial – Versión hispanoamericana

Intente imaginarse los temas descritos a continuación y califique de 0 a 10 qué tan claros y reales se los imaginó. Siendo 0 (no hay imagen en absoluto) y 10 (la imagen es clara y vívida como en la vida real). Use la escala para guiarle. Responda tan pronto como pueda, no piense mucho en esto, tome approximadamente 5 segundos por ítem.

Con "imaginar", nos referimos a crear imágenes mentales de escenarios en su cabeza con los 5 sentidos, por ejemplo, lo que puede ver, incluso con los ojos cerrados, o cuando oimos una canción en nuestra mente aunque no está soniendo. No pase mucho tiempo en cada una, solo conteste con su reacción inmediata.

Electronic Supplementary Material

Factor loadings estimated for the best fitting model in Study 1 (file name = Electronic Supplementary Material 1)

Factor loadings estimated for the best fitting model in Study 2 (Shortened version) (file name = Electronic Supplementary Material 2)

References

Andrade, J., May, J., Deeprose, C., Baugh, S., & Ganis, G. (2014). Assessing vividness of mental imagery: The Plymouth Sensory Imagery Questionnaire. *British Journal of Psychology, 105*(4), 547-563. https://doi.org/10.1111/bjop.12050

Arshamian, A., Manko, P., & Majid, A. (2020). Limitations in odour simulation may originate from differential sensory embodiment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1800), 1-11. https://doi.org/10.1098/rstb.2019.0273

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology-General*, *129*(1), 126-145. https://doi.org/10.1037/0096-3445.129.1.126

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality, *Journal of Public Health*, *27*(3), 281-291. https://doi.org/10.1093/pubmed/fdi031

Bywaters, M., Andrade, J., & Turpin, G. (2004). Determinants of the vividness of visual imagery: The effects of delayed recall, stimulus affect and individual differences. *Memory*, *12*(4), 479-488. https://doi.org/10.1080/09658210444000160

Callow, N., Roberts, R., & Fawkes, J. Z. (2006). Effects of dynamic and static imagery on vividness of imagery, skiing performance and confidence. *Journal of Imagery Research in Sport and Physical Activity*, *1*(1), 1-13. https://doi.org/10.2202/1932-0191.1001

Chalmers, P. (2012). MIRT: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1-29. https://doi.org/10.18637/jss.v048.i06

Chaplin, T. M., & Aldao, A. (2013). Gender differences in emotion expression in children: A meta-analytic review. *Psychological Bulletin*, *139*(4), 735-765. https://doi.org/10.1037/a0030737

Dance, C. J., Ward, J., & Simner, J. (2021). What is the link between Mental Imagery and Sensory Sensitivity? Insights from Aphantasia. *Perception*, *50*(9), 757-782. https://doi.org/10.1177/03010066211042186

Dunlop, B. W., Still, S., LoParo, D., Aponte-Rivera V., Johnson B. N., Schneider R. L., Nemeroff, C. B., Mayberg H. S., & Craighead W. E. (2019). Somatic symptoms in treatment-naïve Hispanic and non-Hispanic patients with major depression. *Depression & Anxiety*, *37*(13), 1-10. https://doi.org/10.1002/da.22984

Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press. https://doi.org/10.4324/9781410605269

Eriksson, M., & Boman, E. (2018). Short is beautiful: Dimensionality and measurement invariance in two length of the basic psychological need satisfaction at work scale. *Frontiers in Psychology*, *9*(965),1-13. https://doi.org/10.3389/fpsyg.2018.00965

Field, A. (2005). *Discovering statistics using* SPSS (2nd ed.). Sage.

Holmes, E. A., Mathews, A., Mackintosh, B., & Dalgleish, T. (2008). The causal effect of mental imagery on emotion assessed using picture-word cues. *Emotion*, *8*(3), 395-409. https://doi.org/10.1037/1528-3542.8.3.395

Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, *30*(3), 349-362. https://doi.org/10.1016/j.cpr.2010.01.001

Jungmann, S., Becker, F., & Witthöft, M. (2022). Measuring the vividness of mental images: An adaptation and validation of German Versions of the Vividness of Visual Imagery Questionnaire (VVIQ) and the Plymouth Sensory Imagery Questionnaire (PSI-Q). *Diagnostica*, *68*(3), 125-136. https://doi.org/10.1026/0012-1924/a000291

Kavanagh, D. J., May, J., & Andrade, J. (2009). Tests of the elaborated intrusion theory of craving and desire: Features of alcohol craving during treatment for an alcohol disorder. *British Journal of Clinical Psychology*, *48*(3), 241-254. https://doi.org/10.1348/014466508X387071

Kozhevnikov, M., Motes, M. A., & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science*, *31*(4), 549-557. https://doi.org/10.1080/15326900701399897

Marks, D. F. (1995). New directions for mental imagery research. *Journal of Mental Imagery*, *19*, 153-167.

McKelvie, S. J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: A critical quantitative review and plea for direction. *Journal of Mental Imagery*, *19*(4), 1-106.

McNorgan, C. (2012). A meta-analytic review of multisensory imagery identifies the neural correlates of modality-specific and modality-general imagery. *Frontiers in Human Neuroscience*, *6*(285), 1-14. https://doi.org/10.3389/fnhum.2012.00285

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental imagery: functional mechanisms and clinical applications. *Trends in Cognitive Sciences*, *19*(10), 590-602. https://doi.org/10.1016/j.tics.2015.08.003

Pérez-Fabello, M. J., & Campos, A. (2020). Spanish version of the Plymouth Sensory Imagery Questionnaire. *Frontiers in Psychology*, *11*(916), 1-10. https://doi.org/10.3389/fpsyg.2020.00916

Perry, L., Morgan, J., Reid, F., Brunton, J., O'Brien, A., Luck, A., & Lacey, H. (2002). Screening for symptoms of eating disorders: Reliability of the SCOFF screening tool with written compared to oral delivery. *International Journal of Eating Disorders*, *32*(4), 466-472. https://doi.org/10.1002/eat.10093

Prydz, E. B., & Wadhwa, D. (2019). Classifying countries by income. *World Bank*. https://datatopics.worldbank.org/world-development-indicators/stories/the-classification-of-countries-by-income.html

Reisberg, D., Pearson, D. G., & Kosslyn, S. M. (2003). Intuitions and introspections about imagery: The role of imagery experience in shaping an investigator's theoretical views. *Applied Cognitive Psychology*, *17*(2), 147-160. https://doi.org/10.1002/acp.858

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36. http://www.jstatsoft.org/v48/i02/

RStudio Team (2016). RStudio: Integrated development for R. Rstudio (Version 1.1456) [Computer software]. http://www.rstudio.com/

Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W. R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *Journal of Business Research*, *58*(7), 935-943. https://doi.org/10.1016/j.jbusres.2003.10.007

Sousa, V., & Rojjanasrirat, W. (2010). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*, *17*(2), 268-274. https://doi.org/10.1111/j.1365-2753.2010.01434.x

Tabachnick, B. G., & Fidell, L. S. (2012). *Using Multivariate Statistics* (6th ed.). Person Education.

Tófoli, L. F., Andrade, L. H, & Fortes, S. (2011). Somatization in Latin America: A review of the classification of somatoform disorders, functional syndromes and medically unexplained symptoms. *Brazilian Journal of Psychiatry*, *33*(Supl. 1), s59-s69. https://doi.org/10.1590/S1516-44462011000500006

Velandia-Morales, A., & Rincon, J. C. (2013). Estereotipos y roles de género utilizados en la publicidad transmitida a través de la televisión. *Universitas Psychologica*, *13*(2), 517-527. https://doi.org/10.11144/Javeriana.UPSY13-2.ergu

World Bank (2019). Population, total - Latin America & Caribbean. *World Bank*. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=ZJ