Open Journal of Statistics, 2026, 16(4), 299-327
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Factor Structure of the BVAQ in Mexican Married Adults

José Moral-de la Rubia

School of Psychology, Universidad Autónoma de Nuevo León, Monterrey, Mexico
Email: jose.morald@uanl.edu.mx

How to cite this paper: Moral-de la Rubia, J. (2026) Factor Structure of the BVAQ in Mexican Married Adults. *Open Journal of Statistics*, **16**, 299-327.
https://doi.org/10.4236/ojs.2026.164014

Received: July 19, 2026
Accepted: August 23, 2026
Published: August 26, 2026

Abstract

The Bermond-Vorst Alexithymia Questionnaire (BVAQ) is an important instrument for assessing alexithymia, offering broader content coverage than the widely used Toronto Alexithymia Scale; however, it has not been validated in Mexico. The objectives of this study were to translate the questionnaire from Dutch into Spanish, test its hierarchical factor structure comprising five first-order factors and two second-order factors, and estimate the reliability and convergent validity of its factors. Alternative models included five correlated factors and a bifactor model. A non-probability quota sample of 570 participants was collected in two phases: one from August to November 2018 (n = 365) and another from January to May 2026 (n = 205). The sample consisted of married men and women residing in Monterrey and its metropolitan area. The hierarchical model with 40 items did not show adequate fit, nor did the alternative models. Using Item Response Theory through five unidimensional models and two bifactor models, items were selected, and all reverse-keyed items were eliminated. The hierarchical model with 20 direct-keyed items and the five-factor correlated model showed good and equivalent fit, both outperforming the bifactor model. It is concluded that the correlated-factors model represents a better option than the hierarchical model because of the adequate convergent validity, reliability, and discriminant validity of its factors, whereas the second-order Affective factor showed weak convergent validity and questionable reliability.

Keywords

Alexithymia, Emotion, Psychometrics, Confirmatory Factor Analysis, Item Response Theory

1. Introduction

Alexithymia emerged within research on psychosomatic disorders and somatization as an alternative explanatory model to the Freudian concept of conversion [1]. It is defined as a disturbance in emotional processing characterized by difficulties in identifying, representing, and verbalizing emotions, limited imaginative activity, and a cognitively concrete, externally oriented thinking style [2] [3]. It is associated with a greater tendency to express emotional distress through somatic symptoms [1].

Because it entails a fundamental deficit in emotional processing, alexithymia is considered a risk factor for various psychopathologies, including psychosomatic disorders, post-traumatic stress disorder, addictions, panic attacks, depression, and eating disorders [4]. Although it shares certain socioemotional difficulties with autism, the latter represents a broader alteration affecting the representation of the internal world in general, not solely affective processing [5] [6]. Currently, alexithymia is conceptualized as a personality trait whose etiology involves both environmental and genetic factors [7]-[9].

Assessment approaches include structured interviews [10] [11], projective techniques [12], ability-based tests [13], and Likert-type psychometric scales [14]-[17]. The latter has demonstrated superior psychometric properties and greater practical acceptance [1]. Prominent examples include the 20-item Toronto Alexithymia Scale (TAS-20) [18], the Bermond-Vorst Alexithymia Questionnaire (BVAQ) [16], and the Perth Alexithymia Questionnaire (PAQ) [19]. The TAS-20 and the PAQ primarily assess difficulties in identifying and expressing emotions, although the PAQ differentiates between positive and negative emotions and incorporates externally oriented thinking. The BVAQ extends this coverage by including impoverished fantasy and reduced emotional reactivity, which are considered central aspects of alexithymia [3].

Bermond *et al.* [20] developed the Amsterdam Alexithymia Scale (AAS), consisting of 20 items. Subsequently, Vorst and Bermond [16] revised and expanded this instrument, resulting in the Bermond-Vorst Alexithymia Questionnaire (BVAQ), which comprises 40 items designed to assess five first-order factors: Verbalizing emotions (VER), Identifying emotions (IDE), Analyzing emotions (ANA), Fantasizing (FAN), and Emotionalizing (EMO). Using principal component analyses (PCA) and confirmatory factor analyses (CFA) based on maximum likelihood estimation, the authors found support for a multidimensional structure in which the five first-order factors could be further organized into two second-order dimensions: Cognitive (COG), comprising Verbalizing, Identifying, and Analyzing, and Affective (AFF), comprising Fantasizing and Emotionalizing [16]. Across two student samples from the University of Amsterdam, the proposed five-factor model showed an acceptable fit despite significant χ^2 values, with χ^2/df ranging from 2.20 to 2.46 and RMSEA ranging from 0.042 to 0.058, although some subscales departed from strict unidimensionality. Multi-group CFA further showed that the constrained five-factor model demonstrated acceptable cross-sex equivalence despite the significant χ^2 difference test for strict invariance. The second-order Cognitive factor (COG) showed a high correlation with the TAS-20 (r =

0.80), whereas the second-order Affective factor (AFF) was independent, reflecting the absence of affective components in the TAS-20. Internal consistency coefficients for the total questionnaire and its subscales ranged from 0.75 to 0.85 [16] [17].

The factorial validity of the BVAQ has been examined extensively since its original development. Early validation studies supported the conceptualization of alexithymia as a multidimensional construct composed of five correlated first-order factors: Verbalizing, Identifying, Analyzing, Fantasizing, and Emotionalizing [16] [17]. In the original Dutch, Belgian, and British samples, the proposed five-factor solution was largely replicated, and subsequent higher-order analyses indicated that these dimensions could be organized into two broader second-order factors representing Cognitive and Affective alexithymia [16]. This distinction was later replicated across six language versions and seven populations, providing cross-cultural support for the hierarchical organization of the construct [21].

Nevertheless, validation studies have produced mixed findings regarding the optimal factorial representation of the instrument. Across populations, unidimensional models have generally shown poor fit and have received little empirical support [17] [22]. By contrast, the five-factor correlated structure has been the most consistently replicated solution, although cross-cultural adaptations have frequently required the removal of problematic items, particularly reverse-keyed indicators. In Japan, Kashimura *et al.* [23] removed 10 items (14, 15, 19, 25, 29, 34, 35, 36, 37, and 39) following principal components analysis because of factor loadings below 0.35. The resulting 30-item version reproduced the expected five-factor structure and showed good internal consistency for the total scale ($\alpha = 0.84$), with factor reliabilities ranging from $\alpha = 0.85$ for Fantasizing to $\alpha = 0.57$ for Emotionalizing. Correlations among factors ranged from 0.42 between Identifying and Verbalizing to 0.16 between Analyzing and Emotionalizing, supporting discriminant validity [23].

Similar findings were reported in Brazil. Using a large community sample of 1285 participants, Salles [24] removed items 4 and 10 because of low factor loading and cross-loadings, respectively. Principal axis factoring with Promax rotation reproduced the original five correlated factors. Internal consistency of the resulting 38-item version was good ($\alpha = 0.83$), whereas subscale reliabilities ranged from $\alpha = 0.85$ for Verbalizing to $\alpha = 0.58$ for Emotionalizing. Significant correlations among factors ranged from 0.44 between Verbalizing and Identifying to 0.09 between Analyzing and Emotionalizing. Moreover, factor scores reproduced the expected higher-order Cognitive and Affective dimensions, thereby supporting the hierarchical interpretation of the questionnaire [24].

Evidence from China also suggested the need for item revision. Wang *et al.* [25] eliminated five items (1, 5, 19, 24, and 30) after an exploratory analysis conducted in a sample of 219 university students. Unlike the original structure, the Identifying dimension split into two distinct factors, labeled Successful Identifying and Unsuccessful Identifying, resulting in a six-factor correlated model. When tested in an independent sample of 220 students, both the six-factor correlated model and the corresponding two-factor higher-order solution showed acceptable fit (χ^2/df = 1.53 and 1.51, *RMSEA* = 0.06, *CFI* = 0.87, *TLI* = 0.85, *SRMR* = 0.07). Internal consistency was acceptable for the total questionnaire (α = 0.79), although reliability varied substantially across factors, ranging from α = 0.78 for Verbalizing and Fantasizing to α = 0.48 for Successful Identifying [25].

In Turkey, Bilge [26] found that the original 40-item five-factor model did not achieve satisfactory fit in a community sample of 600 participants, although principal components analysis reproduced the expected factor configuration with acceptable-to-good reliability coefficients (α = 0.72 - 0.81). Consequently, a simplified 20-item version was developed and tested in an independent sample of 766 participants. Both the correlated five-factor model and the hierarchical model demonstrated good fit; however, the correlated-factors solution showed slightly superior fit indices (χ^2/df = 1.43, *RMSEA* = 0.044, *GFI* = 0.90, *CFI* = 0.96, *IFI* = 0.96) compared with the hierarchical model (χ^2/df = 1.50, *RMSEA* = 0.048, *GFI* = 0.89, *CFI* = 0.96, *IFI* = 0.96). Accordingly, the authors recommended use of the short version under either factorial representation, while acknowledging the statistical advantage of the correlated-factors solution [26].

Additional evidence comes from the Netherlands. de Vroege *et al.* [27] evaluated several structural models in a large general-population sample (n = 1386) and a clinical sample (n = 296). Both the first-order five-factor model and the second-order Cognitive-Affective model demonstrated acceptable fit and measurement invariance across samples. However, the Analyzing factor showed substantial associations with both higher-order dimensions, and psychometric analyses indicated that the first-order subscales provided greater measurement precision and interpretive value than the broader higher-order scores [27].

Overall, accumulated psychometric evidence provides little support for a unidimensional representation of the BVAQ. The distinction between Cognitive and Affective alexithymia has been replicated across multiple countries and populations [21] [22] [24] [27], supporting its value as a higher-order framework. However, the factorial structure most consistently reproduced across validation studies is the model comprising five correlated first-order factors [16] [23]-[27]. At the same time, many adaptations have required item removal or modification to achieve acceptable fit, particularly for reverse-keyed items, and several studies have reported slightly better fit and greater psychometric clarity for the correlated five-factor solution than for the hierarchical alternative [24]-[27]. Thus, current evidence suggests that the Cognitive-Affective distinction is best interpreted as a higher-order organization of the five primary dimensions rather than as an alternative structure of alexithymia itself. Furthermore, the methodological approaches used to evaluate the BVAQ have evolved over time. Whereas earlier psychometric studies primarily relied on standard Maximum Likelihood (ML) estimation [16] [17], more recent validation studies have increasingly employed analytic approaches and estimators better suited to the ordinal nature of Likert-type data, including Robust Maximum Likelihood (MLR) and Diagonally Weighted Least Squares (DWLS/WLSMV) [26] [27].

The BVAQ provides broader factorial coverage of alexithymia than the TAS-20 [18] and complements newer multidimensional measures, such as the Perth Alexithymia Questionnaire [19], by explicitly assessing both cognitive and affective components of the construct [16] [21]. However, despite its cross-cultural use and generally adequate psychometric performance, the instrument has not yet been validated in a Mexican population. Therefore, the aims of the present study were: 1) to translate the BVAQ from Dutch into Spanish and culturally adapt it for use in Mexico; 2) to evaluate the factorial validity of the BVAQ by testing its hypothesized hierarchical structure and comparing it with alternative measurement models using WLSMV estimation for ordinal data; 3) to identify a psychometrically optimized item set through Item Response Theory analyses, qualitatively evaluate the conceptual coverage of the retained items relative to the original instrument, and cross-validate the resulting structure in an independent subsample; and 4) to examine the reliability, convergent validity, discriminant validity, and dimensionality of the resulting scores.

2. Method

2.1. Participants

A non-probability quota sampling strategy based on sex (50% women and 50% men) was employed. Inclusion criteria were being married, being between 18 and 50 years of age, and having the ability to complete a self-administered questionnaire. The only exclusion criterion was an incomplete questionnaire, resulting in the exclusion of 12 cases (7 in Phase 1 and 5 in Phase 2).

The final sample consisted of 570 participants from the general population (285 women and 285 men), all residing in Monterrey and its metropolitan area, Nuevo León, Mexico. Age ranged from 18 to 50 years (M = 34.32, SD = 8.84), and length of marriage ranged from 1 to 30 years (M = 10.51, SD = 8.20). The number of children varied between 0 and 5 (M = 1.91, SD = 1.28).

Regarding occupation, 30.5% were office employees, 23.3% were homemakers, 15.4% were manual workers, 11.9% were technical employees, and 7.9% were students; the remaining categories accounted for less than 5% each. The median level of education corresponded to upper secondary education; 39.5% held a bachelor's degree and 10.2% a postgraduate degree. In terms of subjective socioeconomic status, most participants identified as middle-middle class (54.0%), followed by lower-middle class (25.6%) and upper-middle class (17.5%). With respect to religious affiliation, 69.1% identified as Catholic, 14.6% as Christian, and 8.8% as believers without specific affiliation; 6.7% reported being atheist.

Although the sampling was nonprobabilistic, the order of data collection was examined using the Wald-Wolfowitz runs test [28] as a quality-control procedure. The null hypothesis of randomness was supported for both sociodemographic and psychometric variables in the total sample and in the two data collection subsamples, indicating no evidence of systematic ordering or temporal trends in the sequence of observations. This result should not be interpreted as evidence of participant independence or as changing the nonprobability nature of the sampling design.

2.2. Measurement Instruments

The Vorst and Bermond Alexithymia Questionnaire (BVAQ; Vorst and Bermond, 2001) [16] consists of 40 items rated on a five-point Likert scale (1 = Not at all, 2 = Slightly, 3 = Moderately, 4 = Very much, and 5 = Completely). The questionnaire comprises five factors of eight items each: Verbalizing (directly keyed items: 1, 11, 21, and 36; reverse-keyed items: 6, 16, 26, and 31), Fantasizing (directly keyed items: 7, 17, 22, and 32; reverse-keyed items: 2, 12, 27, and 37), Identifying (directly keyed items: 8, 18, 23, and 33; reverse-keyed items: 3, 13, 28, and 38), Emotionalizing (directly keyed: items: 4, 9, 24, and 34; reverse-keyed items: 14, 19, 29, and 39), and Analyzing (directly keyed items: 5, 15, 25, and 35; reverse-keyed items: 10, 20, 30, and 40). The Emotionalizing and Fantasizing factors define the second-order Affective dimension, whereas the Identifying, Analyzing, and Verbalizing factors constitute the second-order cognitive dimension. Reverse-keyed items were recorded before computing factor and total scores, yielding a total score ranging from 40 to 200, with higher scores indicating greater alexithymia. The internal consistency of the total scale and its factors ranges from acceptable to good across different populations [21]. The initial Spanish version obtained through back-translation is presented in the **Appendix**. As a non-commercial instrument made available by the authors for academic research, the BVAQ was translated and administered solely for scientific purposes.

The BVAQ was administered together with other self-report measures: the Relationship Assessment Scale [29], the Balanced Inventory of Desirable Responding [30], and the Toronto Alexithymia Scale [18], as well as a brief sociodemographic questionnaire. These additional instruments were included as part of a broader assessment protocol but were not analyzed in the present study. The questionnaire began with an informed consent form and did not collect personally identifying information (e.g., names, addresses, telephone numbers, or email addresses), thereby ensuring participants' anonymity and confidentiality.

2.3. Procedure

An instrumental study for psychometric validation was conducted using a cross-sectional ex post facto design based on survey methodology. Data collection was carried out in two independent phases. In the first phase (August-November 2018), university students enrolled in research methodology courses distributed and collected paper questionnaires, yielding 365 participants. In the second phase (January-May 2026), the same recruitment procedures were implemented using an online questionnaire administered through Google Forms, yielding an additional 205 participants. The questionnaire was self-administered and required approximately 20 to 30 minutes to complete.

All participants provided informed consent prior to responding. Participation was voluntary and anonymous, and the confidentiality of the collected information was ensured. To minimize the likelihood of duplicate submissions during the online phase, Google Forms was configured to accept a single response per Google account, and completed questionnaires were additionally screened for potential duplicate entries. This configuration preserved participant anonymity because Google Forms does not disclose account identifiers to researchers. No specific procedures were implemented to prevent participation by both spouses. Recruitment through multiple community contacts and anonymous participation made it unlikely that both members of the same couple would be enrolled.

Because the psychometric analyses required complete questionnaires, cases with one or more missing item responses were excluded, resulting in a final dataset with no missing values. Because recruitment was conducted through multiple community contacts and the total number of individuals approached was not recorded, a participation rate could not be calculated.

2.4. Ethical Considerations

Given its cross-sectional, non-experimental design, the use of informed consent procedures, and the administration of anonymous questionnaires without the collection of personally identifying information, the study was considered to involve no more than minimal risk [31]. Consistent with common research practice for minimal-risk anonymous survey studies at the time, formal ethics committee approval was not sought for either the 2018 or the 2026 data collection. All participants provided written informed consent prior to participation. No personally identifying information (e.g., names, addresses, telephone numbers, or email addresses) was collected, thereby ensuring participant anonymity and data confidentiality.

The study was conducted in accordance with the ethical principles of respect for persons, beneficence, nonmaleficence, and justice governing research involving human participants, as outlined in the Ethical Code of the Sociedad Mexicana de Psicología [32], the American Psychological Association's Ethical Principles of Psychologists and Code of Conduct [33], and the World Medical Association's Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Participants [34].

2.5. Data Analysis

For the first objective, the original Dutch version of the BVAQ was translated into Spanish through independent forward and back translations performed by two bilingual linguistic experts. The translated versions were compared with the original instrument, and discrepancies were resolved by consensus to produce a linguistically equivalent Spanish version [35]. Subsequently, the questionnaire was pilot-tested with 15 Spanish-speaking married adults with an educational level equivalent to compulsory secondary education to assess the comprehensibility of the items. Only minor wording adjustments were required to improve clarity and comprehensibility. The resulting Spanish version is presented in the **Appendix**.

For the second objective, a hierarchical model with five first-order factors and two correlated second-order factors (Cognitive and Affective) was initially tested in the total sample ($n = 570$) using confirmatory factor analysis (CFA). Parameters were estimated using Weighted Least Squares with Mean and Variance adjustment (WLSMV), based on polychoric correlations. Model fit was evaluated using χ^2, χ^2/df, *CFI*, *TLI*, *RMSEA*, and *SRMR*. Values of $\chi^2/df \leq 2$, $RMSEA \leq 0.05$, $SRMR \leq 0.06$, and *CFI* and $TLI \geq 0.95$ were considered indicative of good fit.

Subsequently, item selection was conducted using Item Response Theory (IRT) in the first-phase subsample ($n = 365$). Samejima's graded response model was estimated using the mirt package in R [36] with the default logistic link function. Item discrimination parameter (a) and threshold parameters (b) were estimated under the IRT parameterization, assuming a standard normal latent trait distribution (mean = 0, variance = 1) for model identification. Model fit was evaluated using the M_2 statistic with the C_2 correction, and item fit was assessed using the item-fit statistics implemented in mirt. Unidimensional models were estimated for each first-order factor, along with two bifactor models—one for the cognitive domain and one for the affective domain.

Once the items were selected, three models were tested via CFA in the second-phase subsample ($n = 205$): a hierarchical model, a model with five correlated factors, and a bifactor model with five specific factors. All models were estimated using WLSMV. This method employs Diagonally Weighted Least Squares (DWLS) for parameter estimation but incorporates robust corrections to the χ^2 statistic and standard errors. In all cases, estimates were based on the polychoric correlation matrix and the thresholds of the ordinal variables.

Convergent validity was assessed using the Average Variance Extracted (*AVE*), and reliability was evaluated using McDonald's omega coefficient (ω). *AVE* values ≥ 0.50 indicate convergent validity [37], although lower cutoffs may be acceptable depending on the number of indicators in tau-equivalent measurement models: $AVE \geq 0.54$ with two indicators, ≥ 0.44 with three, and ≥ 0.37 with four, assuming $\omega = 0.70$ [38]. Values of $\omega \geq 0.70$ were considered adequate [39]. In the bifactor model, indices of general factor dominance and unidimensionality were also estimated, including explained common variance (*ECV*), percentage of uncontaminated correlations (*PUC*), omega hierarchical (ω_H), construct reliability (*H*), and factor determinacy (*FD*). In the correlated factors model, discriminant validity was assessed using the heterotrait-monotrait ratio (*HTMT*) of correlations between factors [40]. Analyses were conducted using the lavaan [41], semTools [42], semPlot [43], psych [44], and mirt [36] packages in R.

Finally, the conceptual coverage of the psychometrically optimized version was qualitatively evaluated by the author through a comparison of the retained direct-keyed items with the eliminated reverse-keyed items within each factor. The analysis focused on identifying potential changes in the conceptual breadth of each factor following item reduction and was intended to assess conceptual coverage rather than formal content validity.

3. Results

3.1. Sociodemographic Equivalence between the Two Sampling Phases

Equivalence between the two subsamples was examined in terms of sociodemographic composition and scores on the administered scales. Frequencies by sex, occupational status, and religious affiliation did not differ significantly between the two sampling phases according to the chi-square test ($p > 0.05$). Likewise, the distributions of educational level and subjective socioeconomic status were statistically equivalent based on the Mann-Whitney U test ($p > 0.05$). Additionally, mean age, years of marriage, and number of children did not show significant differences between the two subsamples according to bootstrap-based t tests ($p_{boot} > 0.05$).

3.2. Confirmatory Factor Analysis: Hypothetical Hierarchical Model in the Total Sample

The fit of the hypothetical hierarchical model in the total sample was poor in its initial specification, with eight indicators per first-order factor: χ^2 [df = 734] = 3203.85, $p < 0.001$, $\chi^2/df = 4.36$, $CFI = 0.702$, $TLI = 0.683$, $RMSEA = 0.077$, 95% CI [0.074, 0.080], $p < 0.001$ (H_0: $RMSEA = 0.05$), and $SRMR = 0.102$. AVE values ranged from 0.285 to 0.365. Eight reverse-keyed items showed standardized factor loadings below 0.50 (items 36 in VER; 15 in ANA; 7, 17, and 32 in FAN; and 4, 9, and 34 in EMO). Following the predefined CFA revision criterion, these items were removed and the model was re-estimated. However, model fit remained unsatisfactory and AVE values continued to be low (0.346 - 0.425).

A ninth reverse-keyed item was removed using the same criterion of a loading below 0.50 (item 24 in EMO), yet model fit remained poor and AVE values were still low in four factors (ranging from 0.346 to 0.392). Finally, a tenth reverse-keyed item was eliminated (item 22 in FAN), yielding a model in which only $RMSEA$ reached an acceptable level and the AVE of EMO exceeded 0.50, consistent with the previously revised model. In this final 30-item solution, all factor loadings were greater than 0.50. Across all models, reliability values (ω) were greater than 0.70.

It should be noted that similar results were observed for the alternative model with five correlated factors, which showed poor fit both in its initial specification and in the three subsequent revisions, in which items with factor loadings below 0.50 in the hypothetical (hierarchical) model were removed. The bifactor model solution, included as a first alternative, did not converge under the initial specification or in the first two revisions. In the third revision, the iterative estimation process did converge; however, the model yielded poor fit indices, worse than those obtained for the hierarchical and correlated-factors models.

3.3. Item Selection Using Item Response Theory

3.3.1. Unidimensional Models

As an acceptable fit was not achieved, even with alternative models, the analytical strategy was revised. Item selection for each first-order factor was conducted using IRT. The unidimensionality of each factor was assessed using Horn's parallel analysis [45], retaining a single factor in the five final models (**Table 1**).

Table 1. Goodness of fit, number of factors, average variance extracted, and marginal reliability.

Model	PA	M_2	*df*	*p*	M_2/df	*RMSEA*	*LL*	*UL*	*SRMR*	*TLI*	*CFI*	*AVE*	*MR*
VER	1	4.430	5	0.489	0.886	0	0	0.068	0.029	1	1	0.681	0.803
FAN	1	2.519	2	0.284	1.260	0.027	0	0.128	0.036	0.993	0.998	0.716	0.791
IDE	1	1.593	2	0.451	0.797	0	0	0.097	0.039	1	1	0.704	0.759
EMO	1	5.582	2	0.061	2.791	0.070	0	0.142	0.038	0.976	0.992	0.737	0.817
ANA	1	1.890	2	0.389	0.945	0	0	0.102	0.038	1	1	0.698	0.747
COG		47.672	42	0.253	1.135	0.019	0	0.042	0.034	0.995	0.997	0.523G	
AFF		12.339	12	0.419	1.028	0.009	0	0.054	0.026	0.999	1	0.491G	

Note. Model: VER, FAN, IDE, EMO, and ANA = one-factor models, and COG and AFF = bifactor models. PA = number of items retained through Horn's parallel analysis [45], implemented in the psych package in R [44]. M_2 = goodness-of-fit statistic (C_2 type, moment-based) comparing the marginal proportions and bivariate associations among the items. This statistic approximately follows a χ^2 distribution, with degrees of freedom (*df*) defined as the difference between the number of moments used and the number of estimated parameters. *p* = right-tail probability of the χ^2 distribution with *df* degrees of freedom for testing the null hypothesis of equivalence between the implied and observed moments. M_2/df = normed chi-square statistic. *RMSEA* = root mean square error of approximation; *LL* and *UL* = lower and upper limits of the 95% confidence interval for *RMSEA*. *SRMR* = standardized root mean square residual. *TLI* = Tucker-Lewis index. *CFI* = comparative fit index. *AVE* = average variance extracted in the one-factor model or the general factor (G) in the bifactor models. *MR* = marginal reliability.

The unidimensional models were estimated using marginal maximum likelihood with the mirt package in R, jointly considering indicators of model fit, convergent validity, and reliability. Acceptable to good global fit indices were required, along with generally adequate values of *AVE* and marginal reliability. At the item level, factor loadings of at least 0.50 were required, along with moderate to high discrimination parameters, ordered difficulty parameters within expected ranges, and information curves displaying unimodal profiles, preferably providing greater information at the central levels of the latent trait. Items with low information or inadequate functioning were removed. Overall, the final models demonstrated satisfactory performance according to these criteria (**Table 1** and **Table 2**).

Following this analysis conducted in the phase 1 subsample ($n = 365$), 21 items met the prespecified retention criteria in the unidimensional IRT models: items 6, 11R, 16, 26, and 31 for VER; 2, 12, 27, and 37 for FAN; 3, 13, 28, and 38 for IDE; 14, 19, 29, and 39 for EMO; and 10, 20, 30, and 40 for ANA. These items were subsequently evaluated in the bifactor models of the cognitive and affective domains. Item 11R did not satisfy the retention criteria in the bifactor model and was therefore excluded, resulting in the final 20-item version.

Table 2. Item-level evaluation in the unidimensional models (IRT).

Model	Item	Goodness of fit				λ_F	IRT parameters				
		χ^2	*df*	*p*	*RMSEA*		*a*	b_1	b_2	b_3	b_4
VER	6	17.785	27	0.910	0	0.688	1.615	−1.942	−1.036	0.349	1.832
	11R	38.16	35	0.328	0.016	0.501	0.985	−1.956	−0.412	1.088	2.388
	16	29.641	28	0.381	0.013	0.711	1.722	−1.707	−0.571	0.645	1.951
	26	28.074	28	0.461	0.003	0.681	1.582	−1.947	−0.692	0.822	2.137
	31	22.326	21	0.381	0.013	0.824	2.474	−1.144	−0.076	1.185	2.327
FAN	2	16.063	16	0.449	0.004	0.583	1.223	−2.757	−1.256	0.451	1.787
	12	18.340	15	0.245	0.033	0.716	1.744	−1.600	−0.316	1.080	2.312
	27	19.090	15	0.210	0.037	0.818	2.421	−1.344	−0.635	0.489	1.310
	37	30.332	16	0.016	0.066	0.747	1.911	−1.879	−0.928	0.161	1.189
IDE	3	20.169	21	0.511	0	0.707	1.704	−1.044	−0.003	1.038	2.155
	13	22.356	20	0.322	0.018	0.652	1.463	−1.346	−0.017	1.403	2.986
	28	22.382	19	0.266	0.022	0.744	1.895	−1.000	−0.135	0.920	1.932
	38	18.576	18	0.418	0.009	0.711	1.723	−0.777	0.280	1.534	2.936
EMO	14	25.240	21	0.237	0.024	0.666	1.520	−2.130	−1.002	0.251	1.404
	19	20.834	16	0.185	0.029	0.857	2.831	−1.114	−0.222	0.895	1.812
	29	19.612	18	0.355	0.016	0.799	2.262	−1.501	−0.569	0.465	1.442
	39	24.581	22	0.318	0.018	0.624	1.357	−1.677	−0.611	0.761	2.027
ANA	10	28.565	18	0.054	0.040	0.710	1.715	−0.481	0.622	1.666	2.750
	20	22.456	19	0.262	0.022	0.680	1.578	−1.094	0.167	1.456	2.939
	30	19.417	16	0.248	0.024	0.744	1.894	−0.842	0.414	1.718	2.605
	40	21.836	20	0.349	0.016	0.656	1.481	−1.296	−0.241	1.001	2.213

Note. Unidimensional models: VER, FAN, IDE, EMO, and ANA. One-factor model goodness of fit at the item level: χ^2 = goodness-of-fit statistic, *df* = degrees of freedom, *p* = probability value, and *RMSEA* = root mean square error of approximation. λ_F = factor loadings in the one-factor model estimated by the marginal maximum likelihood method. IRT parameters: *a* = discrimination parameter, and b_1, b_2, b_3, b_4 = difficulty parameters for the response categories.

3.3.2. Bifactor Models

Following item selection based on the unidimensional models, two bifactor models were specified: one for the cognitive domain, with VER, IDE, and ANA as specific factors, and another for the affective domain, with FAN and EMO as specific factors. The models were estimated using marginal maximum likelihood with the mirt package in R, jointly evaluating model fit indices, IRT parameters, and indices associated with general factor dominance, reliability, and factorial specificity.

Item 11R (reverse-keyed) showed inadequate functioning in the cognitive domain, with poor fit and insufficient factor loadings on both the general and specific factors; it was therefore removed. Consequently, subsequent analyses were conducted exclusively with the 20 direct-keyed items.

Both bifactor models demonstrated good overall fit, as well as adequate item-level fit (**Table 3**). At the model level, the cognitive domain showed greater dominance of the general factor (*ECV* = 0.548; *PUC* = 0.727) than the affective domain (*ECV* = 0.370; *PUC* = 0.571). Nevertheless, both models showed high reliability of the general factor, with $\omega = 0.908$ and $\omega_H = 0.854$ for the cognitive domain, and $\omega = 0.904$ and $\omega_H = 0.815$ for the affective domain.

Table 3. Item-level evaluation in the bifactor models (IRT).

Item	Fit		Loadings		IRT parameters								*IECV*
	p	*RMSEA*	λ_G	λ_S	a_1	a_2	a_3	a_4	d_1	d_2	d_3	d_4	
						COG							
6	0.711	0	0.430	0.561	1.036	1.351			3.226	1.715	−0.592	−3.026	0.370
16	0.243	0.018	0.465	0.525	1.109	1.252			2.908	0.957	−1.128	−3.312	0.440
26	0.149	0.022	0.427	0.546	1.008	1.291			3.144	1.117	−1.328	−3.421	0.379
31	0.084	0.028	0.553	0.593	1.606	1.723			2.737	0.142	−2.839	−5.531	0.465
3	0.343	0.013	0.408	0.592	0.998		1.448		1.807	0.001	−1.804	−3.723	0.322
13	0.517	0	0.431	0.486	0.965		1.087		1.960	0.010	−2.052	−4.347	0.441
28	0.372	0.011	0.530	0.527	1.358		1.349		1.896	0.234	−1.775	−3.663	0.503
38	0.621	0	0.531	0.476	1.288		1.154		1.321	−0.518	−2.662	−5.044	0.554
10	0.131	0.025	0.589	0.409	1.437			0.998	0.824	−1.080	−2.887	−4.752	0.675
20	0.77	0	0.670	0.210	1.602			0.501	1.757	−0.286	−2.361	−4.749	0.911
30	0.092	0.027	0.661	0.359	1.705			0.927	1.599	−0.790	−3.282	−4.980	0.772
40	0.262	0.017	0.581	0.278	1.293			0.620	1.892	0.350	−1.458	−3.229	0.813
						AFF							
2	0.639	0	0.591	−0.081	1.253	−0.172			3.507	1.310	−0.506	−2.494	0.982
12	0.313	0.015	0.513	0.324	1.097	0.693			2.002	0.301	−1.346	−3.681	0.715
27	0.138	0.024	0.546	0.791	3.366	4.877			5.976	2.099	−1.884	−7.183	0.323
37	0.053	0.031	0.692	0.270	1.760	0.686			3.483	1.561	−0.392	−2.685	0.868
14	0.136	0.024	0.348	0.566	0.792		1.287		3.227	1.517	−0.386	−2.137	0.274
19	0.125	0.026	0.431	0.740	1.421		2.438		3.166	0.620	−2.552	−5.106	0.254
29	0.068	0.03	0.399	0.693	1.133		1.964		3.412	1.284	−1.065	−3.269	0.250
39	0.463	0.004	0.408	0.480	0.895		1.054		2.297	0.829	−1.056	−2.777	0.419

Note. Bifactor models: COG (cognitive domain) and AFF (affective domain). Fit: p = probability value (right tail) for the null hypothesis of item goodness of fit in the bifactor model; *RMSEA* = root mean square error of approximation. Loadings: λ_G = standardized loading on the general factor; λ_S = standardized loading on the specific factor, estimated by marginal maximum likelihood. IRT parameters: a_1 = discrimination parameter for the general factor; a_2, a_3, a_4 = discrimination parameter for the specific factor; d_1, d_2, d_3, d_4 = threshold (difficulty) parameters associated with the transitions between the five response categories. *IECV* = item explained common variance, or proportion of item variance explained by the general factor.

All items exhibited factor loadings greater than 0.50 on at least one of the two defining factors (general or specific). In the cognitive domain, ANA items were predominantly associated with the general factor, VER items were primarily specific, and IDE items showed a mixed pattern. In the affective domain, FAN items tended to depend more on the general factor, whereas EMO items showed greater specificity (Table 3).

In the cognitive domain, the Verbalizing and Identifying factors retained substantial specific variance, whereas Analyzing depended primarily on the general factor, supporting the use of an overall score complemented by specific factor scores. In contrast, the affective domain showed a more clearly bidimensional structure, in which specific scores become more relevant. Fantasizing exhibited moderate specificity, whereas Emotionalizing depended primarily on specific variance and showed high reliability (Table 4).

Table 4. Factor-level indices in the bifactor model (IRT).

Model	Factor	*ECVSS*	*ECVSG*	*ECVGS*	ω	ω_H	*H*	*FD*
COG	a_1	0.548	0.548	0.548	0.908	0.854	0.993	0.973
	a_2	0.576	0.214	0.424	0.776	0.713	0.875	0.847
	a_3	0.542	0.172	0.458	0.797	0.655	0.927	0.848
	a_4	0.213	0.066	0.787	0.797	0.244	0.996	0.793
AFF	a_1	0.370	0.370	0.370	0.904	0.815	0.909	0.949
	a_2	0.590	0.418	0.410	0.791	0.674	0.745	0.799
	a_3	0.727	0.212	0.273	0.926	0.907	0.929	0.889

Note. Bifactor models: COG (cognitive domain) and AFF (affective domain). Factor: a_1 = general factor; a_2, a_3, and a_4 = specific factors (a_2 = Verbalizing, a_3 = Identifying, and a_4 = Analyzing in COG; a_2 = Fantasizing and a_3 = Emotionalizing in AFF). *ECVSS* = Explained Common Variance of the Specific factor by the Specific factor, or proportion of common variance in a specific factor accounted for by the specific factor; *ECVSG* = Explained Common Variance of the Specific factor by the General factor, or proportion of common variance in a specific factor accounted for by the general factor; *ECVGS* = Explained Common Variance of the General factor by the Specific factors, or proportion of variance in the general factor attributable to the specific factors; ω = total composite reliability; ω_H = hierarchical omega coefficient; *H* = construct reliability; *FD* = factor determinacy.

3.3.3. Item-Selection Sequence

The item-selection process comprised four stages: 1) an exploratory CFA revision of the original 40-item version, during which 10 reverse-keyed items with standardized factor loadings below 0.50 were sequentially removed; because this strategy did not achieve acceptable model fit or convergent validity, it was discontinued; 2) unidimensional IRT analyses of the original 40 items (five first-order factors with eight indicators each), yielding a provisional set of 21 items that met the prespecified retention criteria; 3) bifactor IRT analyses of the resulting cognitive and affective domains, in which Item 11R failed to satisfy the retention criteria and was removed; and 4) cross-validation of the resulting 20-item version by confirmatory factor analysis in the independent second-phase subsample.

3.4. Confirmatory Factor Analysis of Three Models with the Selected Items

In the second sampling phase (n = 205), three models were tested using the 20 selected items: hierarchical (Figure 1), bifactor (Figure 2), and correlated-factors models (Figure 3). In both the hierarchical and correlated-factors models, all items exhibited factor loadings greater than 0.50. In the bifactor model, loadings also exceeded 0.50 on at least one of the two defining factors (general or specific). Correlations among factors in the correlated-factors model ranged from 0.250 (between Identifying and Emotionalizing) to 0.711 (between Identifying and Analyzing), with a mean of 0.460. The correlation between the two second-order factors in the hierarchical model was 0.591 (Table 5 and Table 6).

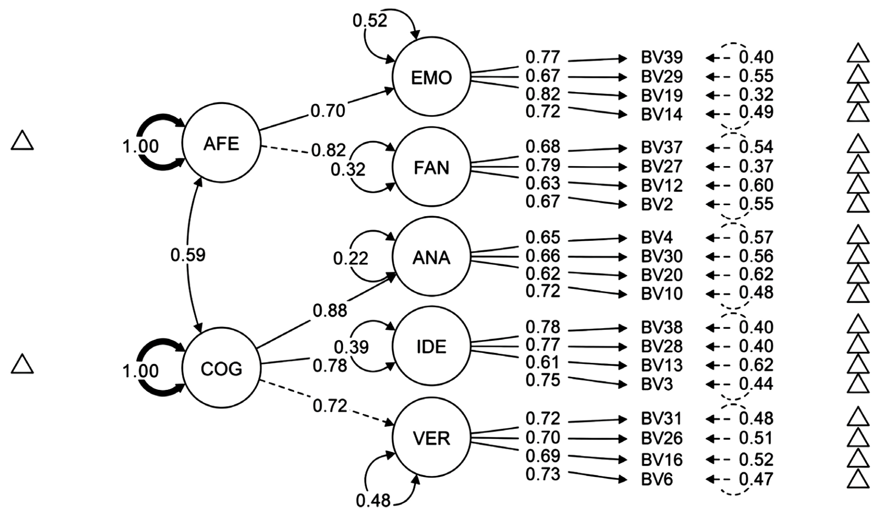

Figure 1. Hierarchical model estimated using WLSMV in the CFA sample (n = 205).

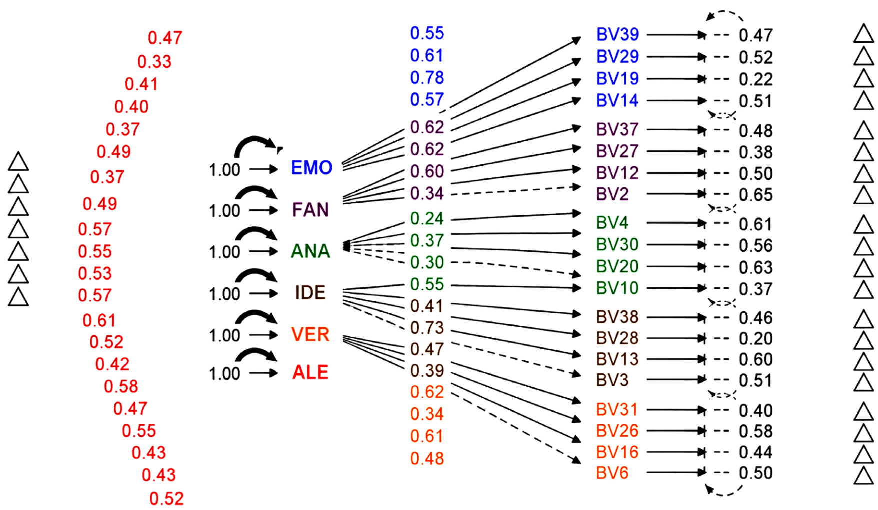

Figure 2. Bifactor model estimated using WLSMV in the CFA sample (n = 205).

Table 5. Factor loadings in three models for the direct-keyed items of the BVAQ and proportion of variance explained by the general factor in the bifactor model.

Factor	Indicator	Hierarchical	Bifactor			Correlated
		λ	λ_G	λ_S	*IECV*	λ
VER→	BV6	0.726	0.521	0.480	0.541	0.728
	BV16	0.694	0.435	0.606	0.340	0.692
	BV26	0.698	0.548	0.345	0.717	0.702
	BV31	0.724	0.465	0.619	0.361	0.720
IDE→	BV3	0.748	0.582	0.394	0.686	0.748
	BV13	0.613	0.417	0.472	0.438	0.615
	BV28	0.773	0.516	0.728	0.335	0.774
	BV38	0.776	0.607	0.413	0.683	0.774
ANA→	BV10	0.721	0.573	0.545	0.524	0.720
	BV20	0.619	0.526	0.302	0.752	0.620
	BV30	0.660	0.546	0.371	0.684	0.660
	BV40	0.653	0.572	0.244	0.846	0.653
FAN→	BV2	0.671	0.490	0.337	0.679	0.670
	BV12	0.633	0.366	0.602	0.271	0.634
	BV27	0.791	0.488	0.620	0.383	0.790
	BV37	0.676	0.374	0.617	0.269	0.677
EMO→	BV14	0.717	0.403	0.571	0.332	0.719
	BV19	0.824	0.413	0.778	0.220	0.823
	BV29	0.673	0.330	0.609	0.227	0.672
	BV39	0.775	0.467	0.554	0.415	0.774
COG→	VER	0.718				
	IDE	0.784				
	ANA	0.882				
AFF→	FAN	0.825				
	EMO	0.695				

Note. Factors: VER = Verbalizing, IDE = Identifying, ANA = Analyzing, FAN = Fantasizing, and EMO = Emotionalizing. Second-order factors: COG = Cognitive and AFF = Affective. Arrow (→) indicates a directional relationship. Items: BV1 - BV40 (reverse-keyed items excluded). Parameters: λ = standardized factor loading; λ_G = loading on the general factor; λ_S = loading on the specific factor. *IECV* = item explained common variance.

Table 6. Correlations between factors in the hierarchical and correlated-factors models.

Model	COG ~~ AFF	VER ~~ IDE	VER ~~ ANA	VER ~~ FAN	VER ~~ EMO	IDE ~~ ANA	IDE ~~ FAN	IDE ~~ EMO	ANA ~~ FAN	ANA ~~ EMO	FAN ~~ EMO
Hierarchical	0.591										
Correlated		0.579	0.589	0.329	0.358	0.711	0.389	0.25	0.443	0.379	0.573

Note. Factors: VER = Verbalizing, IDE = Identifying, ANA = Analyzing, FAN = Fantasizing, Second-order factors: COG = Cognitive and AFF = Affective. ~~ correlation.

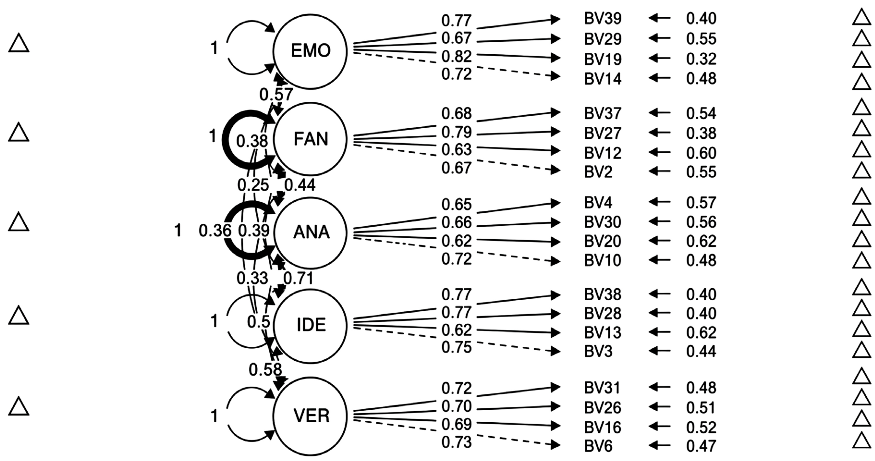

Figure 3. Correlated factors model estimated using WLSMV in the CFA sample (n = 205).

In the bifactor model, Emotionalizing items and most Fantasizing items showed greater specificity than dependence on the general factor (*IECV* < 0.50). Items 16 and 31 of Verbalizing and items 13 and 28 of Identifying were also predominantly specific. In contrast, item 40 of Analyzing showed a clear predominance of the general factor (*IECV* = 0.846) (**Table 5**).

The correlated-factors model showed the best overall fit, followed closely by the hierarchical model, whereas the bifactor model exhibited the least favorable fit (**Table 7**). The hierarchical and correlated-factors models showed statistically equivalent goodness of fit ($\Delta\chi^2 = 6.853$, $\Delta df = 4$, $p = 0.144$), with negligible differences in *CFI*, *TLI*, *RMSEA*, and *SRMR* (all Δ values < 0.01; **Table 8**). By contrast, descriptive comparisons indicated consistently larger differences in fit indices between the bifactor model and the other two models. Accordingly, the hierarchical and correlated-factors models may be regarded as equally plausible representations of the factorial structure of the BVAQ reduced to its direct-keyed items (BVAQ-20).

In the bifactor model, the general alexithymia factor showed high composite reliability ($\omega = 0.922$) and a large proportion of reliable variance attributable to the general factor ($\omega_H = 0.750$), although the explained common variance was moderate (*ECVSS* = 0.459), which does not support a strictly unidimensional structure (**Table 8**). Among the specific factors, Emotionalizing and Fantasizing were the most differentiated from the general factor, whereas Analyzing showed the greatest dependence on it and the lowest specific reliability (ω_H = 0.232; H = 0.428; FD = 0.667). Verbalizing and Identifying exhibited intermediate levels of specificity.

Table 7. Goodness of fit of the three models for the direct-keyed items of the BVAQ.

Statistic or index	Hierarchical		Bifactor		Correlated	
	DWLS	WLSMV	DWLS	WLSMV	DWLS	WLSMV
SCF		1.055		1.107		0.986
SP		71.562		70.560		68.778
χ^2	201.934	262.956	268.721	313.361	188.535	260.045
df	164	164	150	150	160	160
p	0.023	<0.001	<0.001	<0.001	0.061	<0.001
χ^2/df	1.231	1.603	1.791	2.089	1.178	1.625
CFI	0.993	0.947	0.978	0.912	0.995	0.946
TLI	0.992	0.938	0.973	0.888	0.994	0.936
RMSEA	0.034	0.054	0.062	0.073	0.030	0.055
LL	0.013	0.042	0.050	0.062	0	0.043
UL	0.048	0.066	0.074	0.084	0.045	0.067
SRMR	0.061	0.061	0.072	0.072	0.059	0.059

Note. Estimation method: DWLS = diagonally weighted least squares and WLSMV = weighted least squares with mean- and variance-adjusted estimation, based on the polychoric correlation matrix. *SCF* = scaling correction factor and *SP* = shift parameter. χ^2 = chi-square statistic, *df* = degrees of freedom, *p* = probability for the null hypothesis of equivalence between the observed and implied moment matrices; χ^2/df = normed chi-square statistic; *CFI* = comparative fit index; *TLI* = Tucker–Lewis index; *RMSEA* = root mean square error of approximation; *LL* and *UL* = lower and upper limits of the 95% confidence interval for *RMSEA*; *SRMR* = standardized root mean square residual.

Table 8. Comparison of goodness of fit among the three models.

Models	$\lvert\Delta\chi^2\rvert$	$\lvert\Delta df\rvert$	*p*	$\lvert\Delta CFI\rvert$	$\lvert\Delta TLI\rvert$	$\lvert\Delta RMSEA\rvert$	$\lvert\Delta LL\rvert$	$\lvert\Delta UL\rvert$	$\lvert\Delta SRMR\rvert$
H y BF	-	-	-	0.035	0.050	0.019	0.020	0.018	0.011
H y C	6.853	4	0.144	0.001	0.002	0.001	0.001	0.001	0.002
BF y C	-	-	-	0.034	0.048	0.018	0.019	0.017	0.013

Note. Compared models: H = hierarchical, BF = bifactor, and C = correlated factors. $\lvert\Delta\chi^2\rvert$ = absolute difference in chi-square statistics between the compared models (DWLS method with Satorra–Bentler correction); $\lvert\Delta df\rvert$ = degrees of freedom for the goodness-of-fit equivalence test; *p* = probability value for the null hypothesis of goodness-of-fit equivalence; $\lvert\Delta CFI\rvert$ = absolute difference in *CFI*; $\lvert\Delta TLI\rvert$ = absolute difference in *TLI*; $\lvert\Delta RMSEA\rvert$ = absolute difference in *RMSEA*; $\lvert\Delta LL\rvert$ = absolute difference in the lower limits of the 95% confidence interval for RMSEA; $\lvert\Delta UL\rvert$ = absolute difference in the upper limits of the 95% confidence interval for *RMSEA*; $\lvert\Delta SRMR\rvert$ = absolute difference in *SRMR*. $\Delta\chi^2$, Δdf, and *p* values are reported only for the nested comparison between the hierarchical and correlated-factor models. Comparisons involving the bifactor model are descriptive because these models are not nested.

In the hierarchical and correlated-factors models, all factors demonstrated adequate reliability ($\omega > 0.70$), except for the second-order Affective factor in the hierarchical model ($\omega = 0.674$). *AVE* values exceeded 0.50 for most factors; those for Analyzing and Fantasizing were slightly lower, although still compatible with acceptable values for models with four indicators and relatively homogeneous loadings. Finally, in the correlated-factors model, *HTMT* ratios ranged from 0.251 to 0.713, all below 0.90, supporting discriminant validity among factors and the non-unidimensionality of the inventory (**Table 9**).

Table 9. Convergent validity, reliability, and discriminant validity of the factorial models.

Indexes	ALE	VER	IDE	ANA	FAN	EMO	COG	AFF
				Hierarchical model				
AVE		0.505	0.534	0.442	0.483	0.562	0.759	0.713
ω		0.775	0.795	0.724	0.763	0.809	0.808	0.674
				Bifactor model				
ECVSS	0.459	0.529	0.485	0.323	0.622	0.709		
ECVSG	0.459	0.106	0.104	0.056	0.119	0.155		
ECVGS	0.459	0.471	0.515	0.677	0.378	0.291		
ω	0.922	0.808	0.828	0.764	0.793	0.837		
ω_H	0.750	0.420	0.391	0.232	0.488	0.593		
H	0.867	0.621	0.643	0.428	0.659	0.753		
FD	0.885	0.792	0.827	0.667	0.817	0.878		
				Correlated factors model				
AVE		0.505	0.534	0.441	0.483	0.562		
ω		0.803	0.820	0.759	0.788	0.836		
HTMT	VER			0.579	0.590	0.330	0.358	
	IDE				0.713	0.390	0.251	
	ANA					0.443	0.380	
	FAN						0.575	

Note. General factor in the bifactor model: ALE = alexithymia. Specific factors in the bifactor model, first-order factors in the hierarchical model, and singular factors in the correlated-factors model: VER = Verbalizing, IDE = Identifying, ANA = Analyzing, FAN = Fantasizing, and EMO = Emotionalizing. Second-order factors in the hierarchical model: COG = second-order Cognitive factor and AFF = second-order Affective factor. Indices: *AVE* = average variance extracted [37] [38]; ω = McDonald's omega coefficient [39]; *ECVSS* = Explained Common Variance of the Specific factor by the Specific factor, or proportion of variance explained by the specific factor; *ECVSG* = Explained Common Variance of the Specific factor by the General factor, or proportion of variance in the specific factor attributable to the general factor; *ECVGS* = Explained Common Variance of the General factor by the Specific factors, or proportion of variance in the general factor attributable to the specific factors; ω_H = hierarchical omega coefficient; *H* = construct reliability; *FD* = factor determinacy; and *HTMT* = heterotrait-monotrait ratio of correlations proposed by Henseler *et al.* [40].

3.5. Content Coverage Analysis

Because the psychometrically optimized version retained only the direct-keyed items, a qualitative comparison was conducted to examine the extent to which the retained items preserved the conceptual content of the original factors. This analysis was intended to evaluate the conceptual coverage of the retained items rather than to establish formal evidence of content validity.

The comparison indicated that the Identifying and Emotionalizing factors retained essentially the same conceptual content after the removal of the reverse-keyed items, as the retained items represented opposite poles of the same constructs. In contrast, partial reductions in conceptual coverage were observed for the Fantasizing, Verbalizing, and Analyzing factors. In Fantasizing, the retained items mainly assess the production and use of imagination, whereas the discarded reverse-keyed items additionally represent a lack of interest in fantasy and imaginative narratives. In Verbalizing, reverse-keyed items 21 and 36 contribute behavioral and interpersonal aspects that are not explicitly represented by the retained items. In Analyzing, reverse-keyed items 15, 25, and 35 reflect attitudes toward emotional introspection that extend beyond the opposite pole of the retained items. Table 10 summarizes this qualitative comparison.

Table 10. Assessment of conceptual coverage after eliminating the reverse-keyed items.

Factor	Direct-keyed items	Inverse-keyed items
Verbalizing (Cognitive Dimension)	Reverse-keyed items add behavioral and interpersonal aspects not explicitly represented by the retained items	
	6) I like telling others about my emotions. 16) When I want to explain how upset I feel, I can easily find the right words. 26) When something has upset me, I talk to others about my feelings. 31) I can express my feelings in words.	1) I find it difficult to put my feelings into words. 11) Even with a close friend or partner, I find it difficult to talk about my feelings. 21) People often tell me that I should talk more about my feelings. 36) When I talk with people, I tend to talk more about my daily activities than about my feelings.
Fantasizing (Affective Dimension)	Reverse-keyed items broaden the construct by representing lack of interest in fantasy, whereas retained items emphasize imaginative activity.	
	2) Before falling asleep, I often imagine all kinds of events, encounters, and conversations. 12) I often use my imagination. 27) I enjoy inventing strange and imaginative stories. 37) When I have little to do, I daydream.	7) I have few daydreams and fantasies. 17) I have little interest in fairy tales and bizarre stories. 22) I rarely fantasize. 32) I think fantasizing about unreal things or events is a waste of time.
Identifying (Cognitive Dimension)	The retained items appear to preserve the main conceptual domain represented by the original factor.	
	3) When I am upset, I know whether I am anxious, sad, or angry. 13) When things become too much for me, I usually understand why. 28) When I feel bad, I know whether I am afraid, gloomy, or sorrowful. 38) When I am in a sunny mood, I know whether I feel enthusiastic, cheerful, or exuberant.	8) When I am tense, it remains unclear to me which emotion is causing it. 18) When I feel good, it remains unclear to me whether I feel cheerful, elated, or happy. 23) I do not know everything that is going on inside me. 33) When I am struggling with myself, it remains unclear to me whether I am sad, afraid, or unhappy.

Continued

Emotionalizing (Affective Dimension)	The retained items appear to preserve the main conceptual domain represented by the original factor. 14) When friends around me are having a serious argument, I become emotional. 19) Emotions often arise in me spontaneously. 29) Unexpected events often overwhelm me with emotions. 39) When I see someone crying intensely, I feel sadness welling up inside me.	4) When something completely unexpected happens, I remain calm and unmoved. 9) When I see someone crying intensely, I remain unmoved. 24) Even when others are extremely enthusiastic about something, I remain unmoved. 34) I accept disappointments without emotion.
Analyzing (Cognitive Dimension)	Reverse-keyed items represent attitudes toward emotional introspection that are only partially covered by the retained items. 10) One should try to understand one's emotions. 20) When I am not feeling well emotionally, I try to find out why I feel that way. 30) I think people should stay in touch with their feelings. 40) When I am tense, I want to know exactly where that feeling comes from.	5) I hardly ever reflect on my emotions. 15) When I feel bad, I do not further torment myself by wondering why I feel that way. 25) There is little to understand about emotions 35) I find it strange that others analyze their emotions so often.

Note. Author's own elaboration.

Overall, the retained items appeared to preserve the relevance and comprehensibility of the original constructs, although the comprehensiveness of the Fantasizing, Verbalizing, and Analyzing factors was somewhat reduced because certain conceptual facets were no longer represented.

4. Discussion

The first objective of the study was to translate the BVAQ from Dutch into Spanish and culturally adapt it for use in Mexico. As no Spanish version was available at the outset of the study, a back-translation procedure followed by pilot testing to improve item comprehensibility was employed. This approach is considered one of the most rigorous methods for ensuring semantic equivalence across language versions [35]. The final version is presented in the **Appendix**. Although other Spanish translations later became available—such as the one used in Spanish-speaking populations in the United States [22] and another accessible through an artificial intelligence-assisted virtual therapy platform [46]—these emerged after the study began, and in the latter case, the translation lacks academic support.

The second objective was to evaluate the factorial validity of the BVAQ by testing its hypothesized hierarchical structure and comparing it with alternative measurement models. The initial 40-item hierarchical model showed poor fit, even after progressively removing items with standardized factor loadings below 0.50. After item refinement, both the hierarchical and correlated-factors models demonstrated acceptable to good fit and equivalent goodness-of-fit, whereas the bifactor model fit the data significantly worse. Although the general factor exhibited high reliability ($\omega = 0.922$) and accounted for a substantial proportion of reliable variance ($\omega_H = 0.750$), the specific factors retained meaningful variance beyond the general factor, indicating that the structure could not be regarded as essentially unidimensional. Emotionalizing and Fantasizing, corresponding to the affective domain, were the most differentiated from the general factor, whereas Analyzing showed the greatest dependence on it.

The third objective was to identify a psychometrically optimized item set using Item Response Theory (IRT), qualitatively evaluate its conceptual coverage, and cross-validate the resulting structure [47]. IRT analyses conducted in the first-phase subsample supported the elimination of all reverse-keyed items, yielding a balanced 20-item version with four direct-keyed items per factor. The hierarchical and correlated-factors models were subsequently replicated in the independent second-phase subsample, providing evidence for the stability of the abbreviated structure across data collected eight years apart.

These findings should, however, be interpreted in light of the removal of the reverse-keyed items. Although structural validity and reliability improved, some aspects of the original content were not fully retained. A qualitative review of the retained items across factors and dimensions suggested that the conceptual breadth of the Identifying and Emotionalizing factors was largely preserved, whereas the Fantasizing, Verbalizing, and Analyzing factors may have experienced some reduction in conceptual breadth. This trade-off between psychometric optimization and preservation of the original construct representation should be considered when selecting the abbreviated version for research or clinical applications.

The fourth objective was to examine the reliability, convergent validity, discriminant validity, and dimensionality of the resulting scores. The correlated-factors model demonstrated adequate convergent validity (AVE ranging from 0.44 to 0.56), satisfactory reliability (ω ranging from 0.76 to 0.84), and discriminant validity among factors, as assessed through *HTMT*. The hierarchical model yielded comparable results, although the second-order Affective factor showed only questionable reliability ($\omega = 0.67$). The correlation between the second-order Cognitive and Affective factors was moderate ($r = 0.591$), indicating substantial shared variance while supporting the interpretation of alexithymia as a multidimensional construct composed of two related higher-order domains rather than a single general dimension.

These findings are partially consistent with previous BVAQ adaptation studies, all of which required some degree of item reduction to obtain acceptable structural validity. In Brazil, Salles [24] removed two items and replicated the correlated-factors model, although the Emotionalizing factor showed low internal consistency. In China, Wang *et al.* [25] eliminated five items and obtained a six-factor correlated model due to the splitting of the Identifying factor, with some factors showing questionable or unacceptable reliability. In Japan, Kashimura *et al.* [23] removed 10 items and again found weaknesses in the Emotionalizing factor. The Turkish study by Bilge [26] is the closest to the present work, as it reduced the questionnaire to 20 items and obtained good fit for both the hierarchical and correlated-factors models.

Compared with these studies, the present results stand out for the high reliability of the 20-item abbreviated version (ω = 0.92). Moreover, the Emotionalizing factor—which typically exhibits weak psychometric properties when it includes reverse-keyed items—showed the highest indices of convergent validity and reliability in this study, as also reported by Bilge [26]. The Identifying and Verbalizing factors also demonstrated satisfactory properties, whereas Analyzing showed the lowest values, although still within acceptable limits.

Among the study's limitations, the use of non-probability quota sampling should be acknowledged, as it limits the generalizability of the findings and warrants caution when drawing population-level inferences. Although the observed sequence of sociodemographic data was consistent with random ordering according to the Wald-Wolfowitz runs test [28], and the sample exhibited sociodemographic characteristics broadly comparable to those reported for married adults in Monterrey and its metropolitan area [48]—including mean age, years of marriage, number of children, educational attainment, occupational status, subjective socioeconomic status, and religious affiliation—these similarities should not be interpreted as evidence of sample representativeness. Specifically, participants had a mean age of approximately 34 years, had been married for about 10 years, had two children on average, and generally reported upper secondary education. Occupational status was predominantly office or technical employment among men and homemaking or office work among women, while most participants identified as middle socioeconomic status and Catholic. Although cross-validation supported the stability of the abbreviated structure across the two independent subsamples, replication in probabilistic samples and in other Spanish-speaking populations would strengthen the generalizability of the findings. Furthermore, the two sampling phases, conducted eight years apart, displayed comparable sociodemographic and psychometric characteristics, providing additional evidence for the temporal stability and robustness of the proposed measurement model.

The sample size was sufficient for the analyses conducted, meeting minimum recommendations for the number of participants per estimated parameter and overall sample size [49]. The full sample included more than six participants per estimated parameter in the initial hierarchical model (570/86 = 6.6). The first-phase subsample included more than 26 participants per estimated parameter in the unidimensional models (365/14 = 26.1), more than 10 in the bifactor model for the cognitive domain (365/36 = 10.1), and more than 15 in the bifactor model for the affective domain (365/24 = 15.2). The second-phase subsample exceeded 200 participants and included approximately five participants per estimated parameter in the hierarchical model (205/46 = 4.5), with the added advantage of high and relatively homogeneous factor loadings (ranging from 0.613 to 0.882, with a mean of 0.723 in the hierarchical model, and from 0.615 to 0.823, with a mean of 0.708 in the correlated-factors model), supporting the stability of the parameter estimates.

5. Conclusions

This study provides a preliminary Spanish version of the BVAQ for use in Mexico, developed through a translation-back-translation procedure and refined following pilot testing to improve item comprehensibility. The hierarchical model of the original 40-item version did not demonstrate adequate fit, and the reverse-keyed items functioned as weak indicators. After their removal and the reduction of the questionnaire to 20 direct-keyed items, both the hierarchical and correlated-factors models exhibited satisfactory fit and adequate psychometric properties, although the second-order Affective factor showed limited reliability.

The qualitative analysis of conceptual coverage suggested that the reduction to 20 items largely preserved the conceptual content of the Identifying and Emotionalizing factors, whereas some reduction in conceptual breadth may have occurred in the Verbalizing, Fantasizing, and Analyzing factors because the eliminated reverse-keyed items captured additional construct-relevant facets.

Although the overall reliability of the 20-item version was high, the findings do not support a unidimensional structure. The bifactor model showed poorer fit than the hierarchical and correlated-factors models, with greater influence of the general factor within the cognitive domain—particularly for the Analyzing factor—whereas the affective factors retained greater specificity. Overall, the results support the use of the 20-item version as a psychometrically improved measure of alexithymia for research involving Mexican married adults.

6. Recommendations

Considering the limitations associated with non-probability sampling, the use of the abbreviated 20-item version composed of direct-keyed items (BVAQ-20), based on the correlated-factors model, is recommended for married couples in urban Mexican populations, particularly those from Monterrey and its metropolitan area. Its psychometric performance should also be examined in other populations.

Further research should obtain additional evidence of convergent and concurrent validity by comparing the BVAQ-20 with other measures of alexithymia, such as the TAS-20, and with theoretically related constructs. Formal evaluation of content validity and cross-cultural equivalence through expert review and assessment with members of the target population is also recommended. Future studies should examine factorial invariance across sampling periods, sex, and urban-rural contexts, as well as test-retest reliability.

Acknowledgements

The author thanks the reviewers and the editor for their valuable comments, which contributed to improving the manuscript.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Preece, D.A. and Gross, J.J. (2023) Conceptualizing Alexithymia. *Personality and Individual Differences*, **215**, Article 112375. https://doi.org/10.1016/j.paid.2023.112375

[2] Luminet, O. and Nielson, K.A. (2025) Alexithymia: Toward an Experimental, Processual Affective Science with Effective Interventions. *Annual Review of Psychology*, **76**, 741-769. https://doi.org/10.1146/annurev-psych-021424-030718

[3] Sönmez, D., Usta, C.H. and Hocaoğlu, Ç. (2026) Alexisomia in Psychosomatic Medicine: A Review. *Psikiyatride Guncel Yaklasimlar-Current Approaches in Psychiatry*, 1396-1401.

[4] Pinna, F., Manchia, M., Paribello, P. and Carpiniello, B. (2020) The Impact of Alexithymia on Treatment Response in Psychiatric Disorders: A Systematic Review. *Frontiers in Psychiatry*, **11**, Article ID: 311. https://doi.org/10.3389/fpsyt.2020.00311

[5] Ferguson, C.J., Preece, D.A. and Schweitzer, R.D. (2023) Alexithymia in Autism Spectrum Disorder. *Australian Psychologist*, **58**, 131-137. https://doi.org/10.1080/00050067.2023.2174409

[6] Ziaei, T., Sayad Kameh Alia, M., Khoori, E., Shakouri, H., Tatari, M. and Arab, S. (2023) The Comparison of Alexithymia and Self-Consciousness in Male and Female Adolescents. *Current Psychology*, **43**, 15952-15958. https://doi.org/10.1007/s12144-023-05573-2

[7] Davis, S., Rose, S.E., Herron, D. and Elliott, J.M. (2025) Alexithymia: A Narrative Approach to Understanding an Alexithymic Profile. *Journal of Investigative Psychology and Offender Profiling*, **22**, e70002. https://doi.org/10.1002/jip.70002

[8] Ditzer, J., Wong, E.Y., Modi, R.N., Behnke, M., Gross, J.J. and Talmon, A. (2023) Child Maltreatment and Alexithymia: A Meta-Analytic Review. *Psychological Bulletin*, **149**, 311-329. https://doi.org/10.1037/bul0000391

[9] Hernández-Díaz, Y., Genis-Mendoza, A.D., González-Castro, T.B., Fresán, A., Tovilla-Zárate, C.A., López-Narváez, M.L., *et al.* (2024) Exploring Candidate Gene Studies and Alexithymia: A Systematic Review. *Genes*, **15**, Article 1025. https://doi.org/10.3390/genes15081025

[10] Komaki, G., Baba, T., Yoshida, T., Arimura, T., Moriguchi, Y. and Maeda, M. (2024) Development of a Structured Interview for the Modified Version of the Beth Israel Hospital Psychosomatic Questionnaire for Alexithymia. *Frontiers in Psychiatry*, **15**, Article ID: 1356643. https://doi.org/10.3389/fpsyt.2024.1356643

[11] Muzi, S., Di Trani, M., Renzi, A. and Pace, C.S. (2023) Can Alexithymia Be Assessed through an Interview in Adolescents? The Toronto Structured Interview for Alexithymia: Reliability, Concurrent Validity, Discriminant Validity, and Relationships with Emotional-Behavioral Symptoms. *Frontiers in Psychiatry*, **13**, Article ID: 1055946. https://doi.org/10.3389/fpsyt.2022.1055946

[12] Porcelli, P. (2025) Embodied Mentalization—The Case of Human Movement in the Rorschach. *Rorschachiana*, **46**, 3-23. https://doi.org/10.1027/1192-5604/a000189

[13] Lane, R.D. (2023) Measurement of Levels of Emotional Awareness Before, During, and after Psychotherapy. *Journal of Psychotherapy Integration*, **33**, 141-159. https://doi.org/10.1037/int0000297

[14] Preece, D.A., Mehta, A., Petrova, K., Sikka, P., Bjureberg, J., Chen, W., *et al.* (2023) The Perth Alexithymia Questionnaire-Short Form (PAQ-S): A 6-Item Measure of Alexithymia. *Journal of Affective Disorders*, **325**, 493-501. https://doi.org/10.1016/j.jad.2023.01.036

[15] Schroeders, U., Kubera, F. and Gnambs, T. (2022) The Structure of the Toronto Alex-

ithymia Scale (TAS-20): A Meta-Analytic Confirmatory Factor Analysis. *Assessment*, **29**, 1806-1823. https://doi.org/10.1177/10731911211033894

[16] Vorst, H.C.M. and Bermond, B. (2001) Validity and Reliability of the Bermond-Vorst Alexithymia Questionnaire. *Personality and Individual Differences*, **30**, 413-434. https://doi.org/10.1016/s0191-8869(00)00033-7

[17] Zech, E., Luminet, O., Rimé, B. and Wagner, H. (1999) Alexithymia and Its Measurement: Confirmatory Factor Analyses of the 20-Item Toronto Alexithymia Scale and the Bermond-Vorst Alexithymia Questionnaire. *European Journal of Personality*, **13**, 511-532. https://doi.org/10.1002/(sici)1099-0984(199911/12)13:6<511::aid-per347>3.0.co;2-0

[18] Bagby, R.M., Parker, J.D.A. and Taylor, G.J. (1994) The Twenty-Item Toronto Alexithymia Scale—I. Item Selection and Cross-Validation of the Factor Structure. *Journal of Psychosomatic Research*, **38**, 23-32. https://doi.org/10.1016/0022-3999(94)90005-1

[19] Preece, D., Becerra, R., Robinson, K., Dandy, J. and Allan, A. (2018) The Psychometric Assessment of Alexithymia: Development and Validation of the Perth Alexithymia Questionnaire. *Personality and Individual Differences*, **132**, 32-44. https://doi.org/10.1016/j.paid.2018.05.011

[20] Bermond, B., Vorst, H.C.M., Vingerhoets, A.J.J.M. and Gerritsen, W. (1999) The Amsterdam Alexithymia Scale: Its Psychometric Values and Correlations with Other Personality Traits. *Psychotherapy and Psychosomatics*, **68**, 241-251. https://doi.org/10.1159/000012340

[21] Bermond, B., Clayton, K., Liberova, A., Luminet, O., Maruszewski, T., Ricci Bitti, P.E., *et al.* (2007) A Cognitive and an Affective Dimension of Alexithymia in Six Languages and Seven Populations. *Cognition and Emotion*, **21**, 1125-1136. https://doi.org/10.1080/02699930601056989

[22] Morera, O.F., Culhane, S.E., Watson, P.J. and Skewes, M.C. (2005) Assessing the Reliability and Validity of the Bermond-Vorst Alexithymia Questionnaire among U.S. Anglo and U.S. Hispanic Samples. *Journal of Psychosomatic Research*, **58**, 289-298. https://doi.org/10.1016/j.jpsychores.2004.09.001

[23] Kashimura, M., Ogawa, T., Vorst, H.C.M. and Bermond, B. (2011) Psychometric Properties of the Bermond-Vorst Alexithymia Questionnaire in Japanese. *Japanese Psychological Research*, **53**, 302-311. https://doi.org/10.1111/j.1468-5884.2011.00472.x

[24] Salles, B.M. (2023) Bridging the Gap between Alexithymia and Socio-Emotional Impairments: Towards an Intervention Proposal. Doctoral Dissertation, Pontifícia Universidade Católica do Rio de Janeiro. https://www.maxwell.vrac.puc-rio.br/colecao.php?strSecao=resultado&nrSeq=1747&idi=2

[25] Wang, Z., Wang, T., Goerlich, K.S., Pitliya, R.J., Bermond, B., Aleman, A., *et al.* (2021) Psychometric Properties of the Chinese Bermond-Vorst Alexithymia Questionnaire: An Exploratory Structural Equation Modeling Study. *Journal of Pacific Rim Psychology*, **15**, 1-11.

[26] Bilge, Y. (2021) Reliability and Validity of Adaptation to the Turkish of the Bermond-Vorst Alexithymia Questionnaire and Developing of Turkish Short Form. *Journal of Psychiatric Nursing*, **12**, 146-155. https://doi.org/10.14744/phd.2020.79847

[27] de Vroege, L., Emons, W.H.M., Sijtsma, K. and van der Feltz-Cornelis, C.M. (2018) Psychometric Properties of the Bermond-Vorst Alexithymia Questionnaire (BVAQ) in the General Population and a Clinical Population. *Frontiers in Psychiatry*, **9**, Arti-

cle ID: 111. https://doi.org/10.3389/fpsyt.2018.00111

[28] Wald, A. and Wolfowitz, J. (1943) An Exact Test for Randomness in the Non-Parametric Case Based on Serial Correlation. *The Annals of Mathematical Statistics*, **14**, 378-388. https://doi.org/10.1214/aoms/1177731358

[29] Hendrick, S.S. (1988) A Generic Measure of Relationship Satisfaction. *Journal of Marriage and the Family*, **50**, 93-98. https://doi.org/10.2307/352430

[30] Paulhus, D.L. (1998) Manual for Balanced Inventory of Desirable Responding: Version 7 (BIDR-7). Multi-Health Systems. https://search.gesis.org/publication/zis-Paulhus1998Paulhus

[31] Kambhampati, S.B.S., Menon, J. and Maini, L. (2023) Ethics in Research and Publications. *Indian Journal of Orthopaedics*, **57**, 1722-1734. https://doi.org/10.1007/s43465-023-00971-x

[32] Sociedad Mexicana de Psicología (2010) Código Ético del Psicólogo. 5a Edición, Editorial Trillas.

[33] American Psychological Association (2017) Ethical Principles of Psychologists and Code of Conduct. https://www.apa.org/ethics/code/index.aspx

[34] World Medical Association (2024) Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Participants. https://www.wma.net/policies-post/wma-declaration-of-helsinki/

[35] Mumcu, A.Y. (2025) A Review of Translation Processes in Scale Adaptation Studies: A Guide for Scale Translation. *Business Economics and Management Research Journal*, **8**, 119-136. https://doi.org/10.58308/bemarej.1670402

[36] Chalmers, R.P. (2012) Mirt: A Multidimensional Item Response Theory Package for The*r*environment. *Journal of Statistical Software*, **48**, 1-29. https://doi.org/10.18637/jss.v048.i06

[37] Fornell, C. and Larcker, D.F. (1981) Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics. *Journal of Marketing Research*, **18**, 382-388. https://doi.org/10.1177/002224378101800313

[38] Moral de la Rubia, J. (2019) Revisión de los criterios para validez convergente estimada a través de la Varianza Media Extraída. *Psychologia*, **13**, 25-41. https://doi.org/10.21500/19002386.4119

[39] McDonald, R.P. (1999) Test Theory: A Unified Treatment. Lawrence Erlbaum.

[40] Henseler, J., Ringle, C.M. and Sarstedt, M. (2015) A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling. *Journal of the Academy of Marketing Science*, **43**, 115-135. https://doi.org/10.1007/s11747-014-0403-8

[41] Rosseel, Y., Jorgensen, T.D. and De Wilde, L. (2025) LAVAAN: Latent Variable Analysis. Version 0.6-20, Computer Software, Comprehensive R Archive Network.

[42] Jorgensen, T.D., Pornprasertmanit, S., Schoemann, A.M. and Rosseel, Y. (2026) SemTools: Useful Tools for Structural Equation Modeling. Version 0.5-8, Computer Software, Comprehensive R Archive Network.

[43] Epskamp, S. (2025) SemPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output. Version 1.1.7, Computer Software, Comprehensive R Archive Network.

[44] Revelle, W. (2025) Psych: Procedures for Psychological, Psychometric, and Personality Research. Version 2.5.6, Computer Software, Comprehensive R Archive Network.

[45] Horn, J.L. (1965) A Rationale and Test for the Number of Factors in Factor Analysis.

Psychometrika, **30**, 179-185. https://doi.org/10.1007/bf02289447

[46] Freudly (2025) Test psicológicos: Cuestionario de alexitimia de bermond-vorst, BVAQ. https://freudly.ai/es/tests/cuestionario-de-alexitimia-de-bermond-vorst-bvaq/

[47] Hori, K., Fukuhara, H. and Yamada, T. (2022) Item Response Theory and Its Applications in Educational Measurement Part I: Item Response Theory and Its Implementation in R. *WIREs Computational Statistics*, **14**, e1531. https://doi.org/10.1002/wics.1531

[48] INEGI (2025) Estadística de Matrimonios (EMAT). https://www.inegi.org.mx/programas/emat/

[49] Chomeya, R., Tayraukham, S., Tongkhambanchong, S. and Saravitee, N. (2024) Sample Size Determination Techniques for Multivariate Behavioral Sciences Research Emphasizing SEM. *Journal of Education and Innovation*, **26**, 438-447. https://so06.tci-thaijo.org/index.php/edujournal_nu/article/view/270176

Appendix

BVAQ translated from Dutch to Spanish

Señale con una X qué tanto le describe cada una de las siguientes frases en relación su modo habitual de ser, pensar o sentir. Conteste lo más sinceramente posible.

T Totalmente	M Mucho	A Algo	P Poco	N Nada

En relación con modo habitual de ser, pensar o sentir	Me describe
1) Me resulta difícil expresar mis sentimientos con palabras	T M A P N
2) Antes de dormirme, a menudo se me ocurre todo tipo de eventos, encuentros y conversaciones	T M A P N
3) Cuando estoy alterado/a, sé si estoy asustado/a, triste o enfadado/a	T M A P N
4) Cuando ocurre algo completamente inesperado, permanezco tranquilo/a e impasible	T M A P N
5) Casi nunca reflexiono sobre mis emociones	T M A P N
6) Me gusta hablar con otras personas sobre mis sentimientos	T M A P N
7) Tengo pocas ensoñaciones y fantasías	T M A P N
8) Cuando estoy tenso/a, no tengo claro qué emoción lo causa	T M A P N
9) Cuando veo que alguien llora desconsoladamente, permanezco impasible	T M A P N
10) Se debe intentar comprender las emociones	T M A P N
11) Incluso con un amigo, me resulta difícil hablar de mis sentimientos	T M A P N
12) Uso mi imaginación a menudo	T M A P N
13) Cuando las cosas me sobrepasan, generalmente entiendo la causa	T M A P N
14) Cuando los amigos a mi alrededor discuten acaloradamente, me emociono	T M A P N
15) Cuando me siento mal, no me mortifico preguntándome por qué	T M A P N
16) Si quiero decirle a alguien lo mal que me siento, puedo encontrar fácilmente las palabras adecuadas para ello	T M A P N
17) Tengo poco interés en los cuentos de hadas e historias extrañas	T M A P N
18) Cuando me siento bien, no tengo claro si estoy de buen humor, entusiasmado/a o feliz	T M A P N
19) A menudo, las emociones surgen espontáneamente en mí	T M A P N
20) Cuando no me siento bien, intento averiguar por qué me siento así	T M A P N
21) La gente suele decir que debería hablar más sobre mis sentimientos	T M A P N
22) Casi nunca fantaseo	T M A P N
23) No sé lo que está pasando dentro de mí	T M A P N

Continued

24) Incluso cuando otros están muy entusiasmados con algo, permanezco impasible	T	M	A	P	N
25) Hay poco que comprender en las emociones	T	M	A	P	N
26) Cuando algo me altera, hablo con otras personas sobre mis sentimientos	T	M	A	P	N
27) Me gusta inventar historias locas e imaginativas	T	M	A	P	N
28) Cuando me siento mal, sé si estoy asustado/a, decaído/a o triste	T	M	A	P	N
29) Los acontecimientos inesperados a menudo me desbordan de emociones	T	M	A	P	N
30) Creo que hay que mantenerse en contacto con los propios sentimientos	T	M	A	P	N
31) Puedo expresar con palabras mis sentimientos	T	M	A	P	N
32) Considero que fantasear sobre cosas o acontecimientos irreales es una pérdida de tiempo	T	M	A	P	N
33) Cuando me bloqueo emocionalmente, no tengo claro si estoy triste, asustado/a o infeliz	T	M	A	P	N
34) Acepto los contratiempos sin emoción	T	M	A	P	N
35) Me parece extraño que otros analicen sus emociones tan a menudo	T	M	A	P	N
36) Cuando hablo con otras personas, suelo hablar más de mis actividades diarias que de mis sentimientos	T	M	A	P	N
37) Cuando tengo poco que hacer, suelo soñar despierto	T	M	A	P	N
38) Cuando estoy de buen ánimo, sé si estoy entusiasmado/a, alegre o eufórico/a	T	M	A	P	N
39) Cuando veo a otra persona llorar desconsoladamente, siento que me invade la tristeza	T	M	A	P	N
40) Cuando estoy tenso, quiero saber exactamente de dónde viene ese sentimiento	T	M	A	P	N