

Evidence of Convergent Validity of the BVAQ-20 in Relation to the TAS-20 among Mexican Adults and Its Association with Social Desirability

José Moral-de la Rubia

School of Psychology, Autonomous University of Nuevo León, Monterrey, Mexico
Email: jose.morald@uanl.edu.mx

How to cite this paper: Moral-de la Rubia, J. (2026). Evidence of Convergent Validity of the BVAQ-20 in Relation to the TAS-20 among Mexican Adults and Its Association with Social Desirability. *Psychology*, 17, 941-967.

<https://doi.org/10.4236/psych.2026.1710047>

Received: September 6, 2026

Accepted: October 10, 2026

Published: October 13, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In a validation study conducted in Mexico, an abbreviated version of the Bermond-Vorst Alexithymia Questionnaire comprising its 20 direct-keyed items (BVAQ-20) was proposed to achieve good model fit and factors with adequate reliability, convergent validity, and discriminant validity. Although convergent validity between the BVAQ-40 and the Toronto Alexithymia Scale (TAS-20) has been documented in several countries, it has not yet been examined in Mexico using the reduced BVAQ-20 version. Likewise, the extent to which this association is influenced by social desirability remains unknown. The objectives of this study were: (1) to examine the relationship between the BVAQ-20 and the TAS-20 while accounting for social desirability, and (2) to validate the factorial models of the TAS-20 and the control variable. An incidental sample of 570 married Mexican adults (285 women and 285 men) was recruited. Participants completed the BVAQ-40, TAS-20, and the Balanced Inventory of Desirable Responding (BIDR-20). Both the BVAQ-20 and the TAS-20 were correlated with social desirability, although the associations were small. The correlation between the BVAQ-20 and the TAS-20 was high and was primarily accounted for by the Cognitive dimension and its three factors, whereas the Affective dimension and its two factors were largely independent. After accounting for social desirability, the magnitudes of the correlations generally decreased but remained significant. Comparisons of the corresponding partial correlations showed only small differences with trivial effect sizes, providing no consistent evidence that either component of social desirability exerted a meaningfully stronger influence on the associations. These findings provide evidence of the convergent validity of the BVAQ-20 in relation to the TAS-20 and suggest that this association is relatively robust to social desirability.

Keywords

Alexithymia, Convergent Validity, Social Desirability, Self-Deceptive Enhancement, Impression Management, Psychometry

1. Introduction

The psychometric assessment of alexithymia began with the transformation of the Beth-Israel Hospital Psychosomatic Questionnaire (BIQ; [Sifneos, 1973](#)) into a 20-item Likert-type scale known as the Schalling-Sifneos Personality Scale (SSPS; [Apfel & Sifneos, 1979](#); [Sifneos, 1986](#)). However, this instrument exhibited important limitations, including insufficient internal consistency, an unstable factor structure, and incomplete coverage of the construct ([Bagby et al., 1986](#)). In order to overcome these shortcomings, [Taylor et al. \(1985\)](#) developed the Toronto Alexithymia Scale (TAS-26), comprising 26 items distributed across four factors: Difficulty Identifying Feelings (DIF), Difficulty Describing Feelings (DDF), Reduced Daydreaming (RDD), and Externally Oriented Thinking (EOT). Subsequently, [Bagby et al. \(1994\)](#) revised the instrument to improve internal consistency and reduce its association with social desirability by removing fantasy-related items, resulting in the 20-item version (TAS-20) structured into three factors.

Although the TAS-20 became the most widely used instrument for assessing alexithymia in both clinical and research contexts ([Schroeders et al., 2022](#)), some authors argued that its conceptual coverage was limited, particularly with respect to the affective components of the construct ([Bermond et al., 1999](#); [Preece et al., 2024](#)). In response to these concerns, [Bermond et al. \(1999\)](#) developed the Amsterdam Alexithymia Scale (AAS), later revised and expanded by [Vorst and Bermond \(2001\)](#), resulting in the Bermond-Vorst Alexithymia Questionnaire (BVAQ-40).

The BVAQ-40 was designed to assess both the Cognitive dimension of alexithymia—comprising the factors Identifying, Analyzing, and Verbalizing feelings—and the Affective dimension, consisting of the factors Emotionalizing and Fantasizing. The authors proposed a hierarchical model with five first-order factors grouped into two second-order factors (dimensions). Studies conducted primarily in Europe have reported acceptable to good levels of internal consistency and a stable factor structure ([Bermond et al., 2007](#); [Zech et al., 1999](#)). However, subsequent cross-cultural research has identified difficulties mainly associated with reverse-keyed items and has suggested a slight advantage of the correlated five-factor model over the hierarchical model ([Bilge, 2021](#); [Moral-de la Rubia, 2026a](#); [Wang et al., 2021](#)).

Given the established use of the TAS-20 in alexithymia assessment, evidence of the convergent validity of the BVAQ-40 has primarily been examined through its association with this measure. Findings from samples of university students ([Bilge, 2021](#); [Strate, 2025](#); [Vorst & Bermond, 2001](#)), patients with alcohol dependence ([Korkis, 2011](#)), and the general population ([Salles, 2023](#)) have shown a con-

sistent pattern. The total BVAQ-40 score demonstrates high correlations with the total TAS-20 score, while the Cognitive dimension of the BVAQ-40 exhibits very strong associations with this scale. In contrast, the Affective dimension tends to be independent of or only weakly related to the TAS-20, which is consistent with the absence of equivalent affective content in the latter measure. Furthermore, the Verbalizing, Identifying, and Analyzing factors show the strongest associations with DDF, DIF, and EOT, respectively.

Cross-cultural findings regarding difficulties with reverse-keyed items have also been observed in Mexico. In samples of married adults (365 participants in the exploratory phase using item response theory and 205 in the confirmatory phase, after poor fit of the hypothesized hierarchical model in the combined sample of 570), Moral-de la Rubia (2026a) proposed reducing the BVAQ-40 to 20 direct-keyed items (BVAQ-20) due to reliability problems with the 20 reverse-keyed items, resulting in a version similar to that of Bilge (2021). The reliability of the 20 direct-keyed items comprising the BVAQ-20 was very high ($\omega = .922$). The factor with the highest convergent validity and reliability was Emotionalizing ($AVE = .562$, $\omega = .836$), followed by Identifying ($AVE = .534$, $\omega = .820$) and Verbalizing ($AVE = .505$, $\omega = .803$). The lowest values of average variance extracted and reliability were observed for Fantasizing ($AVE = .483$, $\omega = .788$) and Analyzing ($AVE = .441$, $\omega = .759$). These AVE and ω values correspond to the correlated-factor model and were very similar to those obtained for the first-order factors in the hierarchical model. Both the correlated-factor model ($\chi^2/df = 1.625$, $CFI = .946$, $TLI = .936$, $RMSEA = .055$, 95% CI [.043, .067], $SRMR = .059$) and the hierarchical model ($\chi^2/df = 1.603$, $CFI = .947$, $TLI = .938$, $RMSEA = .054$, 95% CI [.042, .066], $SRMR = .061$) showed good and equivalent fit to the data ($|\Delta\chi^2| = 6.853$, $\Delta df = 4$, $p = .144$) using Weighted Least Squares with Mean and Variance adjustment (WLSMV). The Affective dimension, with only two indicators, showed questionable reliability ($\omega = .674$), although it exhibited good convergent validity ($AVE = .713$) due to heterogeneity in its factor loadings (Moral-de la Rubia, 2026a).

After validating the five-factor model in Mexican samples with a substantial item reduction—which only minimally compromised content validity, particularly for the Identifying and Emotionalizing factors—and confirming its stability through multigroup confirmatory factor analyses across independent sampling occasions and sex (Moral-de la Rubia, 2026b), it remains necessary, as part of the BVAQ-20 validation process, to examine its convergent validity in relation to the TAS-20 and its three factors while controlling for social desirability. This variable was one of the reasons for reducing the TAS-26 by six items and one factor (Bagby et al., 1994), and alexithymia measures have been found to be influenced by it (Gaggero et al., 2024).

It should be noted that the TAS-20 has been validated in Mexican samples. In a sample of 381 Mexican university students (Moral-de la Rubia, 2008), the correlated three-factor model showed fit indices ranging from good ($FD = \chi^2/(n - 1)$

= 0.82, $\chi^2/df = 1.88$, $RMSEA = .05$, $SRMR = .05$, $PGI = .96$, $APGI = .95$, $AGFI = .90$) to acceptable ($GFI = .92$, $CFI = .90$, Bollen's $\Delta = .90$) using Maximum Likelihood (ML). In a sample of 200 married adults, the correlated three-factor model showed acceptable fit using ML ($FD = 1.684$, $\chi^2/df = 2.007$, $SRMR = .071$, $RMSEA = .065$, $PGI = .935$, $APGI = .918$, $CFI = .855$, Bollen's $\Delta = .857$), with acceptable factorial variance properties in both samples (Moral-de la Rubia, 2009).

In the sample of 381 students (Moral-de la Rubia, 2008), the TAS-20 demonstrated good internal consistency ($\alpha = .82$) and six-month test-retest reliability ($r = .71$), as did its first two factors: DDF ($\alpha = .80$, $r = .55$) and DIF ($\alpha = .78$, $r = .61$). Reliability for the third factor, EOT, was low ($\alpha = .53$, $r = .36$). The distribution of the total score followed a normal curve according to the Kolmogorov-Smirnov test ($p = .076$), whereas the distributions of its three factors showed positive skewness.

It is also important to note that the TAS-20 is significantly influenced by social desirability (Gaggero et al., 2024). In a sample of 370 Spanish students, the TAS-20 and its three factors correlated with social desirability measured by the Marlowe-Crowne scale (Marlowe & Crowne, 1961), which was used in the TAS-20 development study (Bagby et al., 1994). The Pearson's correlation coefficient between total scores on both scales was $-.41$ and ranged from $-.40$ to $-.22$ across its three factors (Martínez-Sánchez, 1996). In the Mexican validation study, the TAS-20 and its first two factors (DDF and DIF) showed significant inverse correlations with the MMPI L scale, ranging from $-.28$ to $-.29$ (Spearman's coefficient), whereas the third factor (EOT) remained independent (Spearman's $r = -.07$, $p = .186$).

Among contemporary models of social desirability, Paulhus (2002) distinguishes two related but conceptually distinct components: impression management (IM), defined as the tendency to deliberately present oneself in a favorable social light, and self-deceptive enhancement (SDE), defined as an overly positive self-perception that is experienced as genuine. To assess both components, Paulhus developed the Balanced Inventory of Desirable Responding (BIDR-40; Paulhus, 1998), which consists of 40 items. The Marlowe-Crowne scale (Marlowe & Crowne, 1961) and the MMPI L scale primarily measure impression management (Paulhus, 2002).

In a validation study of the BIDR-40 conducted with 624 Mexican university students, it was recommended to eliminate reverse-keyed items, reducing the inventory to 20 items (BIDR-20), with 10 items per factor. The fit of the correlated two-factor model estimated using generalized least squares for the BIDR-20 was good ($FD = 0.47$, $\chi^2/df = 1.74$, $RMSEA = .03$, $GFI = .95$, $AGFI = .94$). The correlation between the two factors was $.67$, with shared variance below 45%, indicating discriminant validity. With this 20-item configuration, internal consistency of the scale ($\alpha = .74$) and its factors were acceptable: SDE ($\alpha = .76$, $\omega = .77$) and IM ($\alpha = .71$, $\omega = .67$) (Moral-de la Rubia et al., 2012).

Although convergent validity between the BVAQ-40 and the TAS-20 has been documented across different cultural contexts, it has not been examined in a Mexican population using the reduced BVAQ-20 version. Likewise, it remains unclear

to what extent the association between both instruments is influenced by social desirability and its two specific components. Clarifying this issue would help determine whether the relationship between these measures primarily reflects assessment of the same construct or is partially explained by response biases. Furthermore, a correlated two-factor model for the BIDR-20 (reduced version) was proposed in a university student sample but has not yet been validated in an independent sample. Additionally, validation of the TAS-20 in Mexican samples has relied on estimation methods that assume multivariate normality, such as ML, despite the availability of more appropriate methods for ordinal variables, such as WLSMV. These studies also employed a six-category ordinal response format (three categories per polarity without a midpoint), whereas the present study uses the original five-category ordinal format.

Therefore, the objectives of this study are to analyze the bivariate relationship of the BVAQ-20 with the TAS-20 and with social desirability, as well as to examine the relationship between the BVAQ-20 and the TAS-20 while statistically controlling for social desirability. This variable is assessed using the BIDR-20, which includes the factors of self-deceptive enhancement (SDE) and impression management (IM). In addition, it will be determined whether the associations between the BVAQ-20 and TAS-20 differ depending on the component of social desirability controlled for by formally comparing the corresponding dependent partial correlation coefficients. Low correlations are expected between both alexithymia measures and social desirability (Gaggero et al., 2024), as well as a high and substantial correlation between the total scores of the BVAQ-20 and the TAS-20, not attributable to response bias (Vorst & Bermond, 2001). It is also expected that this relationship will be primarily explained by the Cognitive dimension of the BVAQ-20 and its three constituent factors.

As a complementary objective, the study aims to validate, through confirmatory factor analysis using WLSMV estimation, the correlated three-factor model of the TAS-20 and the correlated two-factor model of the BIDR-20. Additionally, convergent validity, reliability, and discriminant validity of their factors will be estimated in a sample of married Mexican adults.

2. Method

2.1. Participants

A non-probability quota sampling by sex was used. Inclusion criteria were being married, being between 18 and 50 years of age, and having the ability to complete a self-administered questionnaire. The exclusion criterion was an incomplete questionnaire, which led to the exclusion of twelve surveys.

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. Ages ranged from 18 to 50 years, with a mean of 34.321 years (95% BCa CI [33.573, 35.028]) and a sample standard deviation of 8.844.

Years of marriage ranged from 1 to 30, with a mean of 10.511 years (95% BCa

CI [9.858, 11.233]) and a standard deviation of 8.199. The number of children ranged from 0 to 5, with a mean of 1.909 (95% BCa CI [1.805, 2.007]) and a standard deviation of 1.282.

Regarding occupational status, 174 of the 570 participants (30.5%) reported being office employees, 132 (23.3%) homemakers, 88 (15.4%) manual workers, 68 (11.9%) technical employees, 45 (7.9%) students, 28 (4.9%) small business owners, 15 (2.6%) unemployed, 12 (2.1%) medium-sized business owners, 4 (.7%) retired or permanently disabled, and 4 (.7%) large business owners.

The median educational level corresponded to upper secondary education. Of the 570 participants, 225 (39.5%) reported holding a bachelor's degree, 184 (32.3%) had completed upper secondary education, 95 (16.7%) lower secondary education, 58 (10.2%) postgraduate studies, and 8 (1.4%) primary education.

Both the mode and median of subjective social class corresponded to middle-middle class. Of the 570 participants, 308 (54%) reported belonging to a middle-middle class, 146 (25.6%) middle-lower, 100 (17.5%) middle-upper, 13 (2.3%) lower, and 3 (.5%) upper class.

Regarding religious affiliation, 394 participants (69.1%) identified as Catholic, 83 (14.6%) as Christian, 50 (8.8%) as believers in God without formal religious affiliation, 38 (6.7%) as atheist, and 5 (.9%) reported another affiliation (2 Jewish, 2 Muslim, and 1 Buddhist).

2.2. Measures

The Bermond-Vorst Alexithymia Questionnaire (BVAQ-40; [Vorst & Bermond, 2001](#)) was administered using the Spanish translation by [Moral-de la Rubia \(2026a\)](#). The original instrument consists of 40 items rated on a five-point Likert scale and assesses five factors: Emotionalizing, Fantasizing, Identifying, Analyzing, and Verbalizing. Emotionalizing and Fantasizing define the Affective dimension (AFF), whereas Identifying, Analyzing, and Verbalizing comprise the Cognitive dimension (COG). Based on validation studies in the Mexican population ([Moral-de la Rubia, 2026a, 2026b](#)), the present study employed the reduced BVAQ-20 version, consisting exclusively of 20 direct-keyed items: Verbalizing (VER; items 6, 16, 26, 31), Fantasizing (FAN; items 2, 12, 27, 37), Identifying (IDE; items 3, 13, 28, 38), Emotionalizing (EMO; items 14, 19, 29, 39), and Analyzing (ANA; items 10, 20, 30, 40). Scores are obtained by summing item responses, ranging from 4 to 20 per factor and from 20 to 100 for the total scale. Higher scores indicate greater alexithymia.

The 20-item Toronto Alexithymia Scale (TAS-20; [Bagby et al., 1994](#)) was used in its Spanish version by [Moral-de la Rubia \(2008\)](#). This five-point Likert scale (1 = strongly disagree to 5 = strongly agree) includes 20 items with two response options in each polarity and one midpoint. It comprises three correlated factors: Difficulty Identifying Feelings (DIF; items 1, 3, 6, 7, 9, 13, 14), Difficulty Describing Feelings (DDF; items 2, 4, 11, 12, 17), and Externally Oriented Thinking (EOT; items 5, 8, 10, 15, 16, 18, 19, 20). Scores are obtained by summing item responses

after reverse-keyed items 4, 5, 10, 18, and 19 (6 – item value), yielding a range from 20 to 100. Higher scores indicate higher alexithymia.

The 20-item Balanced Inventory of Desirable Responding (BIDR-20; Moral-de la Rubia et al., 2012) was used. The original version developed by Paulhus (1998) contains 40 items, half positively keyed and half negatively keyed, rated on a 7-point Likert scale (1 = not true to 7 = very true). Based on its validation in Mexican samples (Moral-de la Rubia et al., 2012), a reduced version comprising 20 direct-keyed items (BIDR-20) was used. It assesses two factors: Self-Deceptive Enhancement (SDE; items 1 - 10, corresponding to original items 1, 3, 5, 7, 9, 11, 13, 15, 17, and 19) and Impression Management (IM; items 11 - 20, corresponding to original items 22, 24, 26, 28, 30, 32, 34, 36, 38, and 40). Scores range from 10 to 70 for each factor and from 20 to 140 for the total scale. Higher scores indicate a greater tendency toward socially desirable responding.

2.3. Procedure

This was an empirical instrumental validation study with a cross-sectional ex post facto design conducted using a survey. Participants were contacted through students enrolled in quantitative research and social science research courses. In the initial phase (2018), paper questionnaires were distributed and collected by the students ($n = 365$). In the final phase (2026), students provided a link to an online questionnaire hosted on Google Forms, yielding the remaining 205 participants. Completing the questionnaire required approximately 20 - 30 minutes.

Although the sampling was non-probabilistic, the randomness of the sequence of observations was assessed using the runs test (Wald & Wolfowitz, 1943), with the median as the cutoff point, as a quality-control procedure. In the combined sample ($n = 570$), there was no evidence to reject the null hypothesis of randomness for the sociodemographic variables: sex ($z = -.587$, $p = .557$), age ($z = -.588$, $p = .557$), years of marriage ($z = -.753$, $p = .451$), number of children ($z = 1.418$, $p = .156$), education level ($z = 1.024$, $p = .306$), and subjective socioeconomic status ($z = 1.116$, $p = .264$). Similar results were obtained for the psychometric variables: total scores on the BVAQ-20 ($z = .187$, $p = .852$), BIDR-20 ($z = -1.508$, $p = .132$), and TAS-20 ($z = 1.457$, $p = .145$). Comparable findings were observed in the subsamples from each sampling phase, indicating no evidence of systematic ordering or temporal trends in the sequence of observations across the samples.

Comparability of the subsamples with respect to their sociodemographic composition was also supported. Sex distribution ($\chi^2 [1, N = 570] = .008$, $p = .930$), occupational categories ($\chi^2 [9, N = 570] = 4.239$, $p = .816$), and religious affiliation ($\chi^2 [4, N = 570] = 4.006$, $p = .405$) did not differ significantly between the two sampling phases.

Homogeneity of variability across sampling phases was supported by the Fligner-Killeen test (Fligner & Killeen, 1976) for educational level ($\chi^2 [1, N = 570] = 1.130$, $p = .288$) and subjective socioeconomic status ($\chi^2 [1, N = 570] = 1.436$, $p = .231$). In addition, the Mann-Whitney U test (Mann & Whitney, 1947) indicated

no differences in central tendency for educational level ($z = -.708$, $p = .479$) or subjective socioeconomic status ($z = -0.340$, $p = .734$).

The Brown-Forsythe test (Brown & Forsythe, 1974) supported the assumption of equal variances between sampling phases for age ($F[1, 391.16] = .495$, $p = .482$), years of marriage ($F[1, 403.86] = .030$, $p = .862$), and number of children ($F[1, 437.61] = 0.154$, $p = .695$). Consistent with these results, bootstrap t tests assuming equal variances revealed no significant mean differences in age (mean difference (md) = $-.558$, 95% BCa CI [$-2.140, 1.198$], $p_{boot} = .471$), years of marriage ($md = .127$, 95% BCa CI [$-1.249, 1.610$], $p_{boot} = .856$) and number of children ($md = -.043$, 95% BCa CI [$-.286, .186$], $p_{boot} = .923$). None of these variables followed a normal distribution based on the Shapiro-Wilk test (Royston, 1992): age ($W = .956$, $p < .001$), years of marriage ($W = .902$, $p < .001$), or number of children ($W = .923$, $p < .001$).

2.4. Ethical Considerations

Due to its cross-sectional and non-experimental design, use of informed consent, and anonymous questionnaires without identifying information, the study is classified as minimal risk from an ethical standpoint (Kambhampati et al., 2023). All participants provided informed consent. Anonymity was ensured, and contact information for the principal investigator was provided for questions. The study adhered to ethical standards established by the Mexican Psychological Society (Sociedad Mexicana de Psicología, 2010) and the American Psychological Association (2017).

2.5. Data Analysis

Factorial models for the TAS-20 and BIDR-20 were tested using confirmatory factor analysis (CFA). Parameter estimation was conducted with weighted least squares with mean and variance adjustment (WLSMV), which is particularly appropriate for ordinal data (Park, 2023).

Model fit was evaluated using six indices (Sathyanarayana & Mohanasundaram, 2024): chi-square goodness-of-fit test ($p > .05$ good fit), normed chi-square ($\chi^2/df < 2$ good and < 5 acceptable), Comparative Fit Index ($CFI \geq .95$ good and $\geq .90$ acceptable), Tucker-Lewis Index ($TLI \geq .95$ good and $\geq .90$ acceptable), Root Mean Square Error of Approximation ($RMSEA < .05$ good and $< .08$ acceptable), and Standardized Root Mean Square Residual ($SRMR < .06$ good and $< .10$ acceptable). Convergent validity was assessed using average variance extracted (AVE). AVE values $> .50$ reflect convergent validity (Fornell & Larcker, 1981). However, lower AVE cutoffs can be considered depending on the number of factor indicators (k) (with homogeneous weights) for a composite reliability ω of .70: $AVE > .318$ with five indicators and $> .25$ with seven or more indicators (Moral-de la Rubia, 2019). Composite reliability (internal consistency) of each factor was estimated using coefficient ω (McDonald, 1999). Values of $\omega > .70$ indicate acceptable reliability, $> .80$ good, and $> .90$ excellent. Discriminant validity between the factors of each scale

was assessed using the heterotrait-monotrait correlation ratio (*HTMT*) of Henseler et al. (2015). *HTMT* values < 0.90 indicate discriminant validity. Analyses were conducted using JASP version 0.96 (JASP Team, 2026).

The linear association between BVAQ-20 and TAS-20 was examined using Pearson's product-moment correlation, with partial correlation (Fisher, 1924) used to control for social desirability. The distributional assumption underlying the significance test was evaluated using the H test (Royston, 1983). In addition, at the univariate level, normality (Shapiro-Wilk *W* test; Royston, 1992), outliers (Grubbs, 1969; boxplot; Tukey, 1977), skewness (D'Agostino, 1970), and kurtosis (Anscombe & Glynn, 1983) were assessed. These analyses were performed in R version 4.5.1 (R Core Team, 2025) using the MVN (Korkmaz et al., 2026), nortest (Gross & Ligges, 2015), outliers (Komsta, 2022), and moments (Komsta & Novomestky, 2022) packages.

Given the violation of the multivariate normality assumption, bootstrap procedures were applied to estimate correlations (Rousselet et al., 2023). A total of 2,000 resamples were generated, and 95% confidence intervals (CI) were computed using bias-corrected and accelerated (BCa) methods. Statistical significance was assessed using two-sided 95% bootstrap confidence intervals, with effects considered statistically significant when the interval did not include zero.

To examine whether the observed associations between the BVAQ-20 and TAS-20 varied according to which component of social desirability was statistically controlled for, the corresponding partial correlation coefficients were compared. Because the two partial correlations were dependent and the normality assumption was not met, their differences were evaluated using a nonparametric bootstrap procedure based on 2,000 resamples. The 95% bootstrap percentile confidence interval for each difference was used to assess whether the associations differed significantly. Because eight comparisons were conducted for each TAS-20 outcome, the Benjamini-Hochberg false discovery rate (FDR) procedure (Benjamini & Hochberg, 1995) was applied to the corresponding bootstrap *p* values. Effect size was estimated using Cohen's (1988) *q* statistic. These comparisons were interpreted as differences in the magnitude of the adjusted associations rather than as evidence of the unique contribution of either component of social desirability. Analyses were performed using SPSS version 31 (Cheung et al., 2023; IBM, 2025), with a significance level of .05.

3. Results

3.1. Three-Factor Correlated Model for the TAS-20

Figure 1 presents the standardized solution, showing fairly homogeneous loadings across the three factors. Model fit was not supported by the chi-square test, $\chi^2(167, N = 570) = 816.469, p < .001$. The incremental fit indices indicated good fit (*CFI* = .960 and *TLI* = .955), whereas the normed chi-square ($\chi^2/df = 4.889$) and *SRMR* (.072) were within the prespecified acceptable ranges. The *RMSEA* was slightly above the acceptable cutoff (*RMSEA* = .083, 95% CI [.076, .089]); however,

the test of close fit did not reject the null hypothesis of $RMSEA \leq .08$ ($p = .214$). Thus, the CFA results provided mixed but generally favorable evidence for the hypothesized three-factor structure.

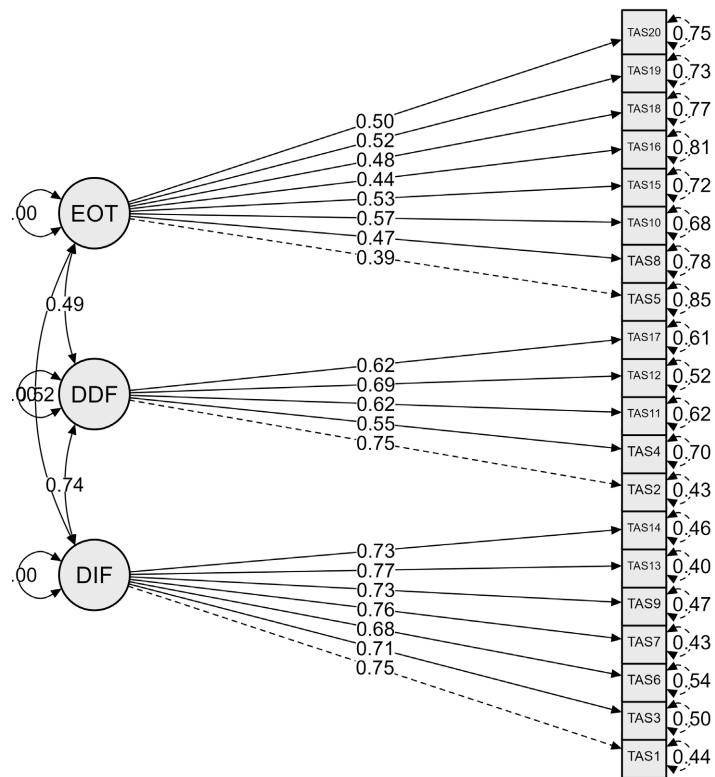


Figure 1. Three-factor correlated model for the TAS-20 estimated using WLSMV.

Overall reliability was high ($\omega = .898 > .80$). DIF showed evidence of convergent validity ($AVE = .538 > .50$) and high reliability ($\omega = .872 > .80$), whereas DDF met the alternative criterion for convergent validity ($AVE = .423 > .32$ with five indicators) and showed acceptable reliability ($\omega = .756 > .70$). In contrast, EOT demonstrated acceptable reliability ($\omega = .711 > .70$), but its AVE (.239) fell slightly below the alternative criterion of .25 for factors with seven or more indicators. Thus, evidence of convergent validity for EOT was weaker than that observed for DIF and DDF. The three factors showed evidence of discriminant validity, with HTMT values of .697 between DIF and DDF, .562 between DIF and EOT, and .444 between DDF and EOT, all below .90.

3.2. Two-Factor Correlated Model for the BIDR-20

Figure 2 presents the standardized solution, showing fairly homogeneous loadings across each of the two factors. Model fit was not supported by the chi-square test, $\chi^2 [169, N = 570] = 595.495, p < .001$. However, the remaining fit indices indicated a fit ranging from good ($SRMR = .049 < .06$) to acceptable ($\chi^2/df = 3.524 < 5$; $CFI = .949$ and $TLI = .942 > .90$; $RMSEA = .067$, 95% CI [.060, .074], $p = 1$ for H_0 : $RMSEA \leq .08$).

Overall reliability was very high ($\omega = .934 > .90$). SDE demonstrated acceptable convergent validity ($AVE = .458 > .25$ for 10 indicators) and high reliability ($\omega = .885 > .80$), as did IM ($AVE = .453 > .25$ with ten indicators and $\omega = .881 > .80$). The two factors showed evidence of discriminant validity ($HTMT = .621 < .90$).

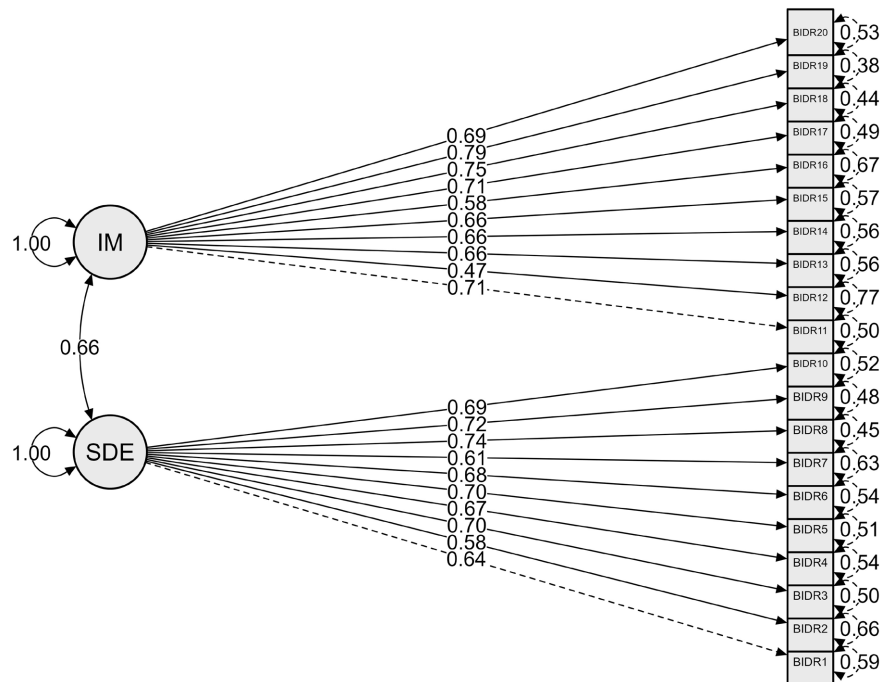


Figure 2. Two-factor correlated model for the BIDR-20 estimated using WLSMV.

3.3. Correlations between the BVAQ-20 and the TAS-20 Controlling for the BIDR-20

3.3.1. Assessment of the Multivariate Normality Assumption

The distributions of the scores for the three scales and all their factors deviated significantly from normality. In the box-and-whisker plots, the BVAQ-20 showed three outliers in each tail and a slight tendency toward negative skewness. The Cognitive dimension of the BVAQ-20, the TAS-20, and the DDF and EOT factors of the TAS-20 presented two outliers in the right tail. The Fantasizing factor of the BVAQ-20 exhibited platykurtosis. The Analyzing factor of the BVAQ-20 and the DIF factor of the TAS-20 showed positive skewness, one outlier in the right tail, and platykurtosis. Finally, the BIDR-20 displayed negative skewness, mainly attributable to the IM factor. For further details, see [Table 1](#).

Consistent with these findings, [Royston's test \(1983\)](#) rejected the null hypothesis of normality for the 68 pairs of variables analyzed, the 96 triplets of variables (BVAQ-20-TAS-20-BIDR-20), and the full set of 15 variables ($H = 183.13$, $p < .001$).

3.3.2. Correlations of the BVAQ-20 and TAS-20 with the BIDR-20

The Fantasizing and Emotionalizing factors, as well as the Affective dimension, were independent of the BIDR-20 total score and its two factors. The Verbalizing

Table 1. Univariate tests for outliers, skewness, kurtosis, and normality.

Var.	Outlier				Skewness			Kurtosis			Normality	
	G	p	I	D	$\sqrt{b_1}$	$z(\sqrt{b_1})$	p	b_2	$z(b_2)$	p	W	p
BVAQ	3.731	<.05	3	3	-.185	-1.812	.070	3.294	1.428	.153	.993	.009
VER	2.483	1			-.064	-.636	.525	2.770	-1.137	.256	.987	<.001
FAN	2.400	1			-.068	-.670	.503	2.465	-3.487	<.001	.985	<.001
IDE	2.882	1			.114	1.125	.261	2.475	-3.393	.001	.980	<.001
EMO	2.231	1			-.049	-.481	.631	2.679	-1.742	.082	.986	<.001
ANA	3.235	.330		1	.235	2.294	.022	2.560	-2.644	.008	.977	<.001
COG	3.411	.174		2	-.101	-.994	.320	2.908	-.327	.744	.993	.008
AFF	2.761	1			-.160	-1.568	.117	2.861	-.592	.554	.994	.019
BIDR	3.590	.088			-.291	-2.814	.005	3.083	.547	.584	.991	.002
SDE	3.598	.085			-.095	-0.937	.349	2.750	-1.262	.207	.991	.001
IM	3.040	.650			-.429	-4.068	<.001	2.770	-1.136	.256	.983	<.001
TAS	2.958	.853		2	-.127	-1.248	.212	2.796	-.976	.329	.994	.023
DIF	2.992	.763		1	.207	2.023	.043	2.427	-3.856	<.001	.978	<.001
DDF	2.750	1		2	-.102	-1.005	.315	2.992	.108	.914	.984	<.001
EOT	3.020	.695		2	-.178	-1.744	.081	3.034	.317	.751	.989	<.001

Note. **Grubbs' (1969)** test: G = test statistic and p = probability value under the null hypothesis of no outlier in the left or right tail; number of outliers in the boxplot: L = left and R = right. **D'Agostino's (1970)** skewness test: $\sqrt{b_1}$ = measure of skewness based on the standardized third central moment, $z(\sqrt{b_1})$ = standardized value of $\sqrt{b_1}$, and p = probability value for a two-tailed test under the null hypothesis of symmetry ($H_0: \sqrt{\beta_1} = 0$). **Anscombe and Glynn's (1983)** kurtosis test: b_2 = measure of kurtosis based on the standardized fourth central moment, $z(b_2)$ = standardized value of b_2 , and p = probability value for a two-tailed test under the null hypothesis of mesokurtosis ($H_0: \beta_2 = 3$). Shapiro-Wilk normality test (**Royston, 1992**): W = test statistic and p = probability value under the null hypothesis of normality.

factor was also independent of the BIDR-20 total score and the IM factor, although it showed a low negative correlation with SDE. In contrast, the BVAQ-20 total score showed low negative correlations with the BIDR-20 total score and the SDE factor. Likewise, the Identifying and Analyzing factors and the Cognitive dimension showed negative correlations with the BIDR-20 total score and its two factors, ranging from $-.110$ to $-.301$ (**Table 2**).

Except for the DDF factor and IM, which were independent, the TAS-20 and its three factors showed low negative correlations with social desirability and its two factors, ranging from $-.104$ to $-.294$. As with the BVAQ-20, the magnitudes of the significant correlations were greater for SDE than for IM (**Table 2**).

3.3.3. Correlations between the BVAQ-20 and the Total Score of the TAS-20 Controlling for the BIDR-20

Correlations between the total score of the TAS-20 and the BVAQ-20 were high

for the overall BVAQ-20 score, Identifying, and the Cognitive dimension; moderate for Analyzing and Verbalizing; and low for Fantasizing and the Affective dimension. In contrast, Emotionalizing was independent of the total TAS-20 score. When controlling for the total score or one of the two factors of the BIDR-20, the significant correlations decreased slightly, with a greater reduction observed when partialling out the SDE factor. Nevertheless, all correlations remained statistically significant (**Table 3**).

When comparing the partial correlations between BVAQ-20 total (or factor) scores and TAS-20 total scores after separately controlling for BIDR-20 Impression Management and Self-Deception, the correlations consistently tended, albeit only slightly, to attenuate less when controlling for Impression Management than

Table 2. Correlations of the Alexithymia Scales with Social Desirability (BIDR-20).

Alexithymia	Social desirability		
	BIDR-20	SDE	IM
BVAQ-20	-.176* [-.275, -.064]	-.240* [-.344, -.125]	-.075 [-.175, .032]
VER	-.066 [-.172, .043]	-.126* [-.239, -.015]	.0028 [-.104, .112]
FAN	.038 [-.069, .150]	-.034 [-.141, .075]	.089 [-.018, .196]
IDE	-.299* [-.386, -.206]	-.290* [-.379, -.194]	-.227* [-.322, -.128]
EMO	.019 [-.094, .132]	.011 [-.105, .128]	.020 [-.083, .125]
ANA	-.224* [-.324, -.124]	-.287* [-.389, -.183]	-.110* [-.206, -.015]
COG	-.251* [-.341, -.151]	-.301* [-.395, -.196]	-.141* [-.237, -.041]
AFF	.037 [-.075, .150]	-.016 [-.133, .101]	.072 [-.033, .179]
TAS-20	-.272* [-.361, -.182]	-.265* [-.360, -.168]	-.206* [-.295, -.115]
DIF	-.294* [-.390, -.195]	-.267* [-.366, -.167]	-.240* [-.331, -.149]
DDF	-.104* [-.208, -.002]	-.132* [-.239, -.028]	-.053 [-.154, .046]
EOT	-.199* [-.291, -.100]	-.193* [-.287, -.093]	-.151* [-.242, -.051]

Note. Pearson's product-moment correlation coefficient with 95% confidence intervals estimated using the BCa method, based on 2000 bootstrap samples and a fixed random seed (2,000,000) to ensure reproducibility. *Indicates that the correlation is significant, as 0 is not included in the interval.

for Self-Deception. Although four of the eight differences were significant at $\alpha = .05$, only two remained significant after Benjamini-Hochberg correction, and both had trivial effect sizes. Therefore, the choice of Impression Management or Self-Deception as the control variable had only a negligible impact on the observed associations (**Table 4**).

Table 3. Correlation Between the BVAQ-20 and the Total Score of the TAS-20 Controlling for the BIDR-20.

BVAQ-20	TAS-20	BIDR-20	SDE	IM
	r [95% BCa CI]	r_p [95% BCa CI]	r_p [95% BCa CI]	r_p [95% BCa CI]
TS	.524* [.448, .599]	.503* [.427, .575]	.492* [.415, .567]	.521* [.444, .597]
VER	.429* [.337, .523]	.428* [.333, .525]	.413* [.321, .507]	.439* [.341, .538]
FAN	.133* [.014, .249]	.149* [.028, .265]	.129* [.010, .247]	.156* [.042, .268]
IDE	.527* [.448, .604]	.486* [.402, .566]	.488* [.405, .568]	.504* [.421, .585]
EMO	.057 [-.045, .160]	.064 [-.038, .164]	.062 [-.040, .163]	.062 [-.043, .164]
ANA	.433* [.345, .522]	.397* [.307, .485]	.387* [.294, .475]	.422* [.333, .512]
COG	.605* [.529, .676]	.576* [.504, .647]	.571* [.496, .647]	.594* [.520, .665]
AFF	.124* [.010, .240]	.140* [.022, .254]	.124* [.008, .243]	.143* [.025, .259]

Note. TS = total score, r = Pearson product-moment correlation, and r_p = partial correlation. 95% BCa CI = 95% confidence interval, estimated by the BCa method from 2,000 bootstrap samples with a seed of 2,000,000 for reproducibility. * indicates a significant correlation, as the confidence interval does not include 0.

Table 4. Comparison of partial correlations between BVAQ-20 and TAS-20 scores after controlling for BIDR-20 Impression Management and Self-Deception.

BVAQ	r_p with TAS		Diff	95% bootstrap CI		Bootstrap		Cohen's q
	IM	SD		LL	UL	p value	BH	
TS	.521	.492	.029	.008	.053	.008*	.013	0.039
VER	.439	.413	.025	.0003	.055	.048	.025	0.031
FAN	.156	.129	.027	-.0001	.057	>.050	.031	0.027
IDE	.504	.488	.016	-.009	.042	.211	.044	0.021
EMO	.062	.062	.0004	-.029	.028	.997	.050	0.0004
ANA	.422	.387	.035	.010	.065	.005*	.006	0.042

Continued

COG	.594	.571	.023	.002	.048	.028	.019	0.035
AFF	.143	.124	.018	-.010	.048	.181	.038	0.019

Note. r_p = partial correlation between the BVAQ-20 total (or factor) score and the TAS-20 total score, controlling separately for Impression Management (IM) and Self-Deception (SD); Diff = difference between the partial correlation controlling for IM and the partial correlation controlling for SD; 95% bootstrap CI = 95% percentile bootstrap confidence interval based on 2000 resamples; *LL* = lower limit; *UL* = upper limit; *p* value = two-tailed bootstrap *p* value (*significant after Benjamini-Hochberg correction); BH = Benjamini-Hochberg false discovery rate correction for a nominal significance level of .05; *q* = Cohen's size effect statistic.

3.3.4. Correlations between the BVAQ-20 and DIF Controlling for the BIDR-20

The correlation of DIF was high with Identifying; moderate with the total BVAQ-20 score and the Cognitive dimension; and low with Verbalizing and Analyzing. In contrast, Fantasizing, Emotionalizing, and the Affective dimension were independent of DIF. When controlling for the total score of the BIDR-20 or either of its two factors, all significant correlations decreased slightly, with the reduction generally being greater when partialling out the SDE factor. Nevertheless, all correlations remained statistically significant (Table 5).

Table 5. Correlation Between the BVAQ-20 and the DIF Factor of the TAS-20 Controlling for the BIDR-20.

BVAQ-20	DIF	BIDR-20	SDE	IM
	r [95% BCa CI]	r_p [95% BCa CI]	r_p [95% BCa CI]	r_p [95% BCa CI]
TS	.346* [.250, .434]	.313* [.223, .399]	.301* [.211, .393]	.339* [.248, .426]
VER	.223* [.110, .334]	.213* [.106, .323]	.198* [.091, .308]	.230* [.120, .343]
FAN	.038 [−.070, .142]	.051 [−.051, .152]	.030 [−.077, .135]	.061 [−.044, .161]
IDE	.506* [.419, .588]	.459* [.369, .541]	.465* [.373, .552]	.478* [.389, .562]
EMO	−.019 [−.130, .095]	−.014 [−.119, .092]	−.017 [−.127, .094]	−.015 [−.123, .094]
ANA	.289* [.193, .387]	.240* [.141, .342]	.231* [.130, .335]	.273* [.178, .371]
COG	.440* [.353, .523]	.396* [.307, .482]	.392* [.302, .483]	.423* [.335, .506]
AFF	.014 [−.096, .120]	.026 [−.075, .127]	.010 [−.094, .114]	.032 [−.072, .136]

Note. TS = total score, r = Pearson product-moment correlation, and r_p = partial correlation. 95% BCa CI = 95% confidence interval, estimated by the BCa method from 2000 bootstrap samples with a seed of 2,000,000 for reproducibility. *Indicates a significant correlation, as the confidence interval does not include 0.

When comparing the partial correlations between the BVAQ-20 total (or factor) score and DIF score after separately controlling for BIDR-20 Impression Management and Self-Deception, correlations tended to attenuate slightly less when controlling for Impression Management than for Self-Deception, except for the non-significant association with Emotionalizing. Six of eight differences were significant at $\alpha = .05$, but only two remained significant after Benjamini-Hochberg correction, and in this case, the effect size was trivial. Therefore, the choice of Impression Management or Self-Deception as the control variable had a minimal effect on the observed associations (Table 6).

Table 6. Comparison of partial correlations between BVAQ-20 and DIF scores controlling for BIDR-20 Impression Management and Self-Deception.

BVAQ	r_p with DIF		Diff	95% bootstrap CI		Bootstrap		Cohem's q
	IM	SD		LL	UL	p value	BH	
TS	.339	.301	.037	.009	.070	.009	.006	0.042
VER	.230	.198	.032	.001	.069	.039	.031	0.034
FAN	.061	.030	.031	.003	.067	.024	.019	0.031
IDE	.478	.465	.013	-.013	.040	.353	.044	0.017
EMO	-.015	-.017	.002	-.029	.032	.906	.050	0.002
ANA	.273	.231	.042	.010	.082	.010*	.013	0.045
COG	.423	.392	.031	.001	.066	.038	.025	0.037
AFF	.032	.010	.022	-.008	.055	.126	.038	0.022

Note. r_p = partial correlation between the BVAQ-20 total (or factor) score and the DIF score, controlling separately for Impression Management (IM) and Self-Deception (SD); Diff = difference between the partial correlation controlling for IM and the partial correlation controlling for SD; 95% bootstrap CI = 95% percentile bootstrap confidence interval based on 2,000 resamples; LL = lower limit; UL = upper limit; p value = two-tailed bootstrap p value (*significant after Benjamini-Hochberg correction); BH = Benjamini-Hochberg false discovery rate correction for a nominal significance level of .05; q = Cohen's size effect statistic.

3.3.5. Correlations between the BVAQ-20 and DDF Controlling for the BIDR-20

Correlations between DDF and Verbalizing and the Cognitive dimension were high, and moderate with the total BVAQ-20 score, Identifying, and Analyzing. In contrast, Fantasizing, Emotionalizing, and the Affective dimension were independent of DDF. When controlling for the total score of the BIDR-20 or either of its two factors, all significant correlations—except for one—decreased slightly, although they remained statistically significant. The reduction was generally greater when partialling out the IM factor. The exception was observed in the correlation between Verbalizing and DDF, which showed a slight increase when controlling for this factor (Table 7).

When comparing the partial correlations between the BVAQ-20 total (or factor) score and DDF score after separately controlling for BIDR-20 Impression

Table 7. Correlation Between the BVAQ-20 and the DDF Factor of the TAS-20 Controlling for the BIDR-20.

BVAQ-20	DDF	BIDR-20	SDE	IM
	r [95% BCa CI]	r_p [95% BCa CI]	r_p [95% BCa CI]	r_p [95% BCa CI]
TS	.468* [.377, .551]	.460* [.368, .542]	.454* [.362, .536]	.466* [.375, .549]
VER	.563* [.474, .646]	.561* [.471, .646]	.556* [.467, .640]	.564* [.476, .648]
FAN	.070 [−.054, .194]	.074 [−.048, .200]	.066 [−.052, .190]	.075 [−.047, .202]
IDE	.362* [.255, .468]	.348* [.240, .457]	.341* [.236, .448]	.360* [.251, .464]
EMO	.098 [−.017, .212]	.100 [−.017, .215]	.100 [−.020, .214]	.099 [−.017, .211]
ANA	.303* [.192, .412]	.289* [.177, .394]	.279* [.167, .388]	.300* [.191, .406]
COG	.543* [.449, .625]	.537* [.446, .619]	.533* [.443, .614]	.542* [.449, .623]
AFF	.106 [−.016, .232]	.110 [−.009, .235]	.105 [−.018, .228]	.110 [−.012, .236]

Note. TS = total score, r = Pearson product-moment correlation, and r_p = partial correlation. 95% BCa CI = 95% confidence interval, estimated by the BCa method from 2,000 bootstrap samples with a seed of 2,000,000 for reproducibility. *Indicates a significant correlation, as the confidence interval does not include 0.

Management and Self-Deception, no differences were significant, although correlations tended to attenuate less when controlling for Impression Management than for Self-Deception, except for the association with Emotionalizing (Table 8).

3.3.6. Correlations between the BVAQ-20 and EOT Controlling for the BIDR-20

Correlations between EOT and the total BVAQ-20 score, as well as the Analyzing, Identifying, and Cognitive dimensions, were moderate, whereas correlations with Verbalizing, Fantasizing, and the Affective dimension were low. In contrast, Emotionalizing was independent of EOT. When controlling for the total score or either of the two BIDR-20 factors, the correlations showed small changes, tending to increase in most cases. Nevertheless, all correlations remained statistically significant (Table 9).

When comparing partial correlations between the BVAQ-20 total (or factor) scores and EOT scores after separately controlling for BIDR-20 Impression Management and Self-Deception, the differences for BVAQ-20 total score, Fantasizing, and Verbalizing were significant, with a trivial effect size, and did not remain significant after Benjamini-Hochberg correction. These three correlations and the five nonsignificant correlations tended to attenuate less when controlling for Impression Management than for Self-Deception (Table 10).

Table 8. Comparison of partial correlations between BVAQ-20 and DDF scores controlling for BIDR-20 Impression Management and Self-Deception.

BVAQ	r_p with DDF		Diff	95% bootstrap CI		Bootstrap		Cohen's q
	IM	SD		LL	UL	p value	BH	
TS	.466	.454	.013	-.005	.032	0.158	0.019	0.016
VER	.564	.556	.008	-.001	.024	0.100	0.006	0.012
FAN	.075	.066	.009	-.007	.028	0.275	0.031	0.009
IDE	.360	.341	.019	-.007	.047	0.188	0.025	0.021
EMO	.099	.100	-.001	-.017	.014	0.852	0.050	-0.001
ANA	.300	.279	.020	-.005	.049	0.108	0.013	0.022
COG	.542	.533	.009	-.012	.033	0.383	0.038	0.013
AFF	.110	.105	.006	-.009	.024	0.455	0.044	0.006

Note. r_p = partial correlation between the BVAQ-20 total (or factor) score and the DDF score, controlling separately for Impression Management (IM) and Self-Deception (SD); Diff = difference between the partial correlation controlling for IM and the partial correlation controlling for SD; 95% bootstrap CI = 95% percentile bootstrap confidence interval based on 2000 resamples; LL = lower limit; UL = upper limit; p value = two-tailed bootstrap p value (*significant after Benjamini-Hochberg correction); BH = Benjamini-Hochberg false discovery rate correction for a nominal significance level of .05; q = Cohen's size effect statistic.

Table 9. Correlation Between the BVAQ-20 and the EOT Factor of the TAS-20 Controlling for the BIDR-20.

BVAQ-20	EOT	BIDR-20	SDE	IM
	r [95% BCa CI]	r_p [95% BCa CI]	r_p [95% BCa CI]	r_p [95% BCa CI]
TS	.445 [.349, .534]	.425 [.336, .512]	.418 [.329, .504]	.440 [.347, .529]
VER	.281 [.167, .390]	.274 [.167, .380]	.264 [.158, .369]	.285 [.178, .391]
FAN	.231 [.125, .330]	.243 [.135, .344]	.228 [.119, .330]	.248 [.142, .345]
IDE	.327 [.225, .427]	.286 [.177, .392]	.288 [.184, .392]	.304 [.200, .407]
EMO	.083 [-.018, .183]	.089 [-.011, .189]	.087 [-.013, .188]	.088 [-.013, .188]
ANA	.441 [.349, .524]	.415 [.322, .499]	.411 [.318, .493]	.432 [.341, .515]
COG	.452 [.357, .541]	.424 [.331, .512]	.421 [.329, .510]	.440 [.347, .530]
AFF	.206 [.102, .304]	.218 [.112, .314]	.207 [.102, .302]	.220 [.118, .316]

Note. TS = total score, r = Pearson product-moment correlation, and r_p = partial correlation. 95% BCa CI = 95% confidence interval, estimated by the BCa method from 2000 bootstrap samples with a seed of 2,000,000 for reproducibility. *Indicates a significant correlation, as the confidence interval does not include 0.

Table 10. Comparison of partial correlations between BVAQ-20 and EOT scores controlling for BIDR-20 Impression Management and Self-Deception.

BVAQ	r_p with EOT		Diff	95% bootstrap CI		Bootstrap		Cohem's q
	IM	SD		LL	UL	p value	BH	
TS	.440	.418	0.021	.002	.045	.032	.006	0.026
VER	.285	.264	0.021	.000	.047	.049	.019	0.023
FAN	.248	.228	0.019	.001	.043	.043	.013	0.021
IDE	.304	.288	0.016	-.014	.046	.303	.044	0.017
EMO	.088	.087	0.000	-.022	.021	.999	.050	0.000
ANA	.432	.411	0.022	.000	.047	.056	.025	0.026
COG	.440	.421	0.019	-.005	.047	.116	.031	0.024
AFF	.220	.207	0.013	-.007	.037	.177	.038	0.014

Note. r_p = partial correlation between the BVAQ-20 total (or factor) score and the EOT score, controlling separately for Impression Management (IM) and Self-Deception (SD); Diff = difference between the partial correlation controlling for IM and the partial correlation controlling for SD; 95% bootstrap CI = 95% percentile bootstrap confidence interval based on 2000 resamples; LL = lower limit; UL = upper limit; p value = two-tailed bootstrap p value (*significant after Benjamini-Hochberg correction); BH = Benjamini-Hochberg false discovery rate correction for a nominal significance level of .05; q = Cohen's size effect statistic.

4. Discussion

Both the BVAQ-20 and the TAS-20 showed significant correlations with social desirability, although these were generally low, with only one reaching a value of $-.301$; thus, the proportion of shared variance was less than one-eleventh. Within the BVAQ-20, the Cognitive dimension and its three factors were related to social desirability, whereas the Affective dimension and its two factors were independent of it. Consistent with this finding, the TAS-20 and its three factors were also related to social desirability, which is consistent with their emphasis on the cognitive aspects of alexithymia (Bagby et al., 1994; Vorst & Bermond, 2001). Consequently, the TAS-20 total score showed a higher correlation with social desirability than the BVAQ-20 total score, as was also observed in a previous study conducted in Novosibirsk, Russia (Dorosheva & Burova, 2025).

The correlations were negative, indicating that higher levels of social desirability were associated with lower reported levels of alexithymia. Alexithymic traits, which imply lower emotional and social intelligence, are not generally considered socially desirable (Gaggero et al., 2024). The pattern of findings provides only limited evidence that the association between alexithymia and social desirability is more strongly linked to Self-Deception than to Impression Management. Although attenuation of correlations was sometimes greater when controlling for Self-Deception, particularly for TAS-20 total scores and DIF, the opposite tendency was observed for DDF, whereas correlations involving EOT often showed

slight increases after statistical control. Moreover, most differences between partial correlations were small, frequently failed to remain significant after Benjamini-Hochberg correction, and were associated with trivial effect sizes. Taken together, these findings suggest that neither component of social desirability exerts a substantively stronger influence on alexithymia scores and that the observed associations are largely robust to the separate control of Self-Deception and Impression Management.

In line with expectations regarding convergent validity of the BVAQ-20 in relation to the TAS-20 (Vorst & Bermond, 2001), the BVAQ-20 total score and its Cognitive dimension showed high bivariate correlations with the TAS-20 total score. Likewise, DIF was most strongly associated with Identifying, DDF with Verbalizing, and EOT with Analyzing, reflecting the conceptual overlap between corresponding components of the two instruments. Correlations between DDF and Verbalizing and between DIF and Identifying were high, whereas those between EOT and Analyzing were moderate. In contrast, the Affective dimension and its two factors showed little or no association with the TAS-20 factors. This pattern is consistent with the theoretical distinction between cognitive and affective components of alexithymia proposed by Vorst and Bermond (2001) and provides further support for the convergent validity of the BVAQ-20.

It should be noted that the correlations observed were slightly lower than those reported among Dutch psychology students by Vorst and Bermond (2001) and in the Brazilian general population by Salles (2023). However, the correlation magnitudes were more similar to those reported by Bilge (2021) among Turkish nursing students, in which the BVAQ-40 had been reduced by 20 items, and by Korkis (2011) among Dutch patients receiving treatment for alcohol dependence.

When the effects of the two components of social desirability were considered separately, the correlations between the BVAQ-20 and TAS-20 measures generally decreased slightly but remained statistically significant. Notably, the correlations between the total scores of the BVAQ-20 and TAS-20 and between Verbalizing and DDF remained high. Across analyses involving TAS-20 total score and its three dimensions (DIF, DDF, and EOT), differences between partial correlations controlling for Impression Management versus Self-Deception were generally small and associated with trivial effect sizes. Thus, the results do not provide evidence that either component of social desirability exerts a substantially stronger influence on the associations between alexithymia measures. The persistence of the correlations after adjustment indicates that the convergent associations between the BVAQ-20 and TAS-20 primarily reflect shared alexithymia-related variance rather than variance attributable to response-style biases.

An exception to the general attenuation pattern was observed for EOT, for which controlling for Impression Management frequently produced slight increases in correlation magnitude. However, these differences were small, associated with trivial effect sizes, and did not remain significant after correction for multiple testing. Overall, the findings indicate that the associations between the

BVAQ-20 and TAS-20 are relatively robust to the separate statistical control of Impression Management and Self-Deception, with neither component showing a meaningful differential influence on the observed relationships.

Regarding the secondary objective of validating the three-factor correlated model of the TAS-20 using an estimation method more appropriate for ordinal items than that employed in previous studies with Mexican samples (i.e., maximum likelihood; Moral-de la Rubia, 2008, 2009), the findings provided mixed but generally favorable evidence for the model. The *CFI* and *TLI* indicated good fit, whereas the χ^2/df and SRMR were within the prespecified acceptable ranges. Although the RMSEA point estimate (.083) slightly exceeded the acceptable cutoff of .08, its 95% confidence interval included values below .08 and the test of close fit did not reject the hypothesis that $RMSEA \leq .08$ ($p = .214$). Reliability was high for the total scale and the DIF factor and acceptable for the DDF and EOT factors (Levant et al., 2024; Salles, 2023; Strate, 2025). Convergent validity, assessed through average variance extracted, was good for DIF and acceptable for DDF, particularly when evaluated according to the criteria proposed by Moral-de la Rubia (2019). In contrast, EOT showed weaker evidence of convergent validity, as its *AVE* (.239) fell slightly below the alternative criterion of .25 for factors with seven or more indicators. This finding is consistent with the relatively weak convergent validity commonly reported for EOT in psychometric studies of the TAS-20 (Bagby et al., 1994; Levant et al., 2024; Salles, 2023; Strate, 2025). This may be attributable to the fact that four of its five items are reverse-keyed and that its content is more heterogeneous than that of the other two factors. In contrast, discriminant validity among the three factors was supported by *HTMT* values below .90. These findings are broadly consistent with previous results obtained in samples of university students (Moral-de la Rubia, 2008) and married couples (Moral-de la Rubia, 2009), in which maximum likelihood estimation was used.

Validation of the two-factor correlated model of the BIDR-20 also yielded acceptable fit indices. Reliability was very high, convergent validity was acceptable, particularly according to the criteria proposed by Moral-de la Rubia (2019), and discriminant validity was good. Consequently, support was found for this reduced two-factor correlated model composed exclusively of direct-keyed items among married Mexican adults, although the model was originally proposed in a different population, namely university students (Moral-de la Rubia et al., 2012).

Several limitations should be acknowledged. First, the sampling procedure was non-probabilistic; therefore, the findings should be interpreted with due caution. Nevertheless, the runs test provided no evidence against the null hypothesis of randomness in the sequence of observations across the data-collection phases. In addition, the two independent subsamples—collected eight years apart and subsequently combined into the total sample—were statistically equivalent with respect to sociodemographic and psychometric variables. Their demographic characteristics were also consistent with those reported for the population of married adults in Monterrey and its metropolitan area (INEGI, 2025). The average partic-

ipant was approximately 34 years old, had been married for about 10 years, had two children, and had completed upper-secondary education. Occupational roles showed a gendered pattern: office or technical employment predominated among men, whereas homemaking and office employment predominated among women. Additionally, most participants reported a middle subjective socioeconomic status and identified as Catholic. Finally, the sample size was large ($n = 570 > 500$), which—combined with the use of bias-corrected and accelerated (BCa) confidence intervals in the presence of multivariate non-normality—enhanced the robustness of significance testing (Rousselet et al., 2023).

A further limitation is that the BVAQ-40 was administered in its original format rather than in the 20-item version composed of direct-keyed items analyzed in the present study (BVAQ-20). Consequently, these findings require replication using the final version of the scale.

5. Conclusion

The findings provide evidence of the convergent validity of the BVAQ-20 with respect to the TAS-20 in a sample of married Mexican adults. As expected given the conceptual foundations of both instruments, the strongest associations were observed between the Cognitive dimension of the BVAQ-20 and the TAS-20, whereas the Affective dimension and its factors showed weak or negligible relationships with the TAS-20 dimensions. Moreover, the pattern of associations among the factors of the two instruments was consistent with the constructs they were designed to assess, supporting the distinction between cognitive and affective components of alexithymia.

Both the BVAQ-20 and the TAS-20 showed statistically significant, albeit generally small, associations with social desirability. Although controlling separately for Impression Management and Self-Deception resulted in slight changes in the correlations between the two alexithymia measures, the evidence of convergent validity remained largely unchanged. Comparisons of the corresponding partial correlations revealed only small differences, typically associated with trivial effect sizes, and provided no consistent evidence that either component of social desirability exerted a meaningfully stronger influence on the observed associations. Taken together, these findings indicate that the convergent associations between the BVAQ-20 and TAS-20 are relatively robust to the separate statistical control of social desirability and primarily reflect shared alexithymia-related variance rather than response-style bias.

The three-factor correlated model of the TAS-20 received mixed but generally favorable support in the present sample. The model showed good CFI and TLI and acceptable χ^2/df and SRMR values, although the RMSEA point estimate slightly exceeded the prespecified acceptable cutoff of .08. Reliability was high for the total scale and DIF and acceptable for DDF and EOT. Convergent validity was good for DIF and acceptable for DDF according to the criteria proposed by Moral-de la Rubia (2019), whereas EOT showed weaker evidence of convergent validity

because its AVE fell slightly below the alternative criterion. The three factors showed evidence of discriminant validity. Overall, these findings support the use of the TAS-20 three-factor structure in this sample, while indicating that the EOT factor warrants particular caution and further investigation.

6. Future Research

Future research should examine the convergent validity of the BVAQ-20 with respect to the Perth Alexithymia Questionnaire (PAQ; Preece et al., 2018), as well as its concurrent validity with measures of depression (Kieraité et al., 2024) and emotional distress (Bagby et al., 2026; Preece et al., 2024). In addition, the simplified BVAQ-20 version should be administered in these studies.

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

- American Psychological Association (2017). *Ethical Principles of Psychologists and Code of Conduct*. <https://www.apa.org/ethics/code/index.aspx>
- Anscombe, F. J., & Glynn, W. J. (1983). Distribution of the Kurtosis Statistic b_2 for Normal Samples. *Biometrika*, 70, 227-234. <https://doi.org/10.1093/biomet/70.1.227>
- Apfel, R. J., & Sifneos, P. E. (1979). Alexithymia: Concept and Measurement. *Psychotherapy and Psychosomatics*, 32, 180-190. <https://doi.org/10.1159/000287386>
- Bagby, R. M., Mortezaei, A., Lau, S. C. L., McIntyre, C. S., & Taylor, G. J. (2026). Discriminant Validity of the 20-Item Toronto Alexithymia Scale and the Perth Alexithymia Questionnaire in Relation to Psychological Distress. *Behavioral Sciences*, 16, Article 339. <https://doi.org/10.3390/bs16030339>
- Bagby, R. M., Parker, J. D. A., & 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](https://doi.org/10.1016/0022-3999(94)90005-1)
- Bagby, R. M., Taylor, G. J., & Ryan, D. P. (1986). The Measurement of Alexithymia: Psychometric Properties of the Schalling-Sifneos Personality Scale. *Comprehensive Psychiatry*, 27, 287-294. [https://doi.org/10.1016/0010-440x\(86\)90004-0](https://doi.org/10.1016/0010-440x(86)90004-0)
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 57, 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- 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>
- Bermond, B., Vorst, H. C. M., Vingerhoets, A. J. J. M., & Gerritsen, W. (1999). The Am-

- sterdam Alexithymia Scale: Its Psychometric Values and Correlations with Other Personality Traits. *Psychotherapy and Psychosomatics*, 68, 241-251.
<https://doi.org/10.1159/000012340>
- 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>
- Brown, M. B., & Forsythe, A. B. (1974). The Small Sample Behavior of Some Statistics Which Test the Equality of Several Means. *Technometrics*, 16, 129-132.
<https://doi.org/10.1080/00401706.1974.10489158>
- Cheung, S. F., Pesigan, I. J. A., & Vong, W. N. (2023). DIY Bootstrapping: Getting the Nonparametric Bootstrap Confidence Interval in SPSS for Any Statistics or Function of Statistics (When This Bootstrapping Is Appropriate). *Behavior Research Methods*, 55, 474-490. <https://doi.org/10.3758/s13428-022-01808-5>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge.
- D'Agostino, R. B. (1970). Transformation to Normality of the Null Distribution of g_1 . *Biometrika*, 57, 679-681. <https://doi.org/10.2307/2334794>
- Dorosheva, E. A., & Burova, T. S. (2025). Psikhometricheskie kharakteristiki russkoyazychnoy versii oprosnika aleksitimii B. Bermonda i G. Vorsta (BVAQ). *Sibirskiy Psikhologicheskiy Zhurnal*, 96, 66-82.
- Fisher, R. A. (1924). The Distribution of the Partial Correlation Coefficient. *Metron*, 3, 329-332.
- Fligner, M. A., & Killeen, T. J. (1976). Distribution-Free Two-Sample Tests for Scale. *Journal of the American Statistical Association*, 71, 210-213.
<https://doi.org/10.1080/01621459.1976.10481517>
- Fornell, C., & 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>
- Gaggero, G., Balboni, G., & Esposito, G. (2024). The Role of Cultural Capital in Self-Reported Alexithymia and Empathy. *Psicothema*, 36, 267-276.
<https://doi.org/10.7334/psicothema2023.372>
- Gross, J., & Ligges, U. (2015). *Nortest: Tests for Normality* (Version 1.0-4) [Computer Software]. Comprehensive R Archive Network.
<https://CRAN.R-project.org/package=nortest>
- Grubbs, F. E. (1969). Procedures for Detecting Outlying Observations in Samples. *Technometrics*, 11, 1-21. <https://doi.org/10.1080/00401706.1969.10490657>
- Henseler, J., Ringle, C. M., & 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>
- IBM (2025). *IBM SPSS Statistics for Windows (Versión 31)* [Computer Software].
<https://www.ibm.com/mx-es/products/spss-statistics>
- INEGI (2025). *Estadística de Matrimonios (EMAT)*.
<https://www.inegi.org.mx/programas/emat/>
- JASP Team (2026). *JASP* (Versión 0.96) [Computer Software]. <https://jasp-stats.org/>
- Kambhampati, S. B. S., Menon, J., & Maini, L. (2023). Ethics in Research and Publications. *Indian Journal of Orthopaedics*, 57, 1722-1734.
<https://doi.org/10.1007/s43465-023-00971-x>

- Kieraité, M., Bättig, J. J., Novoselac, A., Noboa, V., Seifritz, E., Rufer, M. et al. (2024). "Our Similarities Are Different": The Relationship between Alexithymia and Depression. *Psychiatry Research*, 340, Article 116099. <https://doi.org/10.1016/j.psychres.2024.116099>
- Komsta, L. (2022). *Outliers: Tests for Outliers* (Version 0.15) [Computer Software]. Comprehensive R Archive Network. <https://CRAN.R-project.org/package=outliers>
- Komsta, L., & Novomestky, F. (2022). *Moments: Moments, Cumulants, Skewness, Kurtosis and Related Tests* (Version 0.14.1) [Computer Software]. Comprehensive R Archive Network. <https://CRAN.R-project.org/package=moments>
- Korkis, L. (2011). *Constructvaliditeit van de Bermond-Vorst Alexithymia Questionnaire (BVAQ-40) bij alcoholafhankelijke patiënten*. Master's Thesis, Universiteit Twente. <https://essay.utwente.nl/essays/61626>
- Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2026). *MVN: Multivariate Normality Tests* (Version 6.3) [Computer Software]. Comprehensive R Archive Network. <https://CRAN.R-project.org/package=MVN>
- Levant, R. F., Parent, M. C., Ling, S., & Borgogna, N. C. (2024). Variance Composition, Reliability, Validity, and Measurement Invariance of the Toronto Alexithymia Scale-20 in Non-Clinical Samples 1. In *Assessing and Treating Emotionally Inexpressive Men* (pp. 61-80). Routledge. <https://doi.org/10.4324/9781003378518-6>
- Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18, 50-60. <https://doi.org/10.1214/aoms/1177730491>
- Marlowe, D., & Crowne, D. P. (1961). Social Desirability and Response to Perceived Situational Demands. *Journal of Consulting Psychology*, 25, 109-115. <https://doi.org/10.1037/h0041627>
- Martínez-Sánchez, F. (1996). Adaptación española de la Escala de Alexitimia de Toronto (TAS-20). *Clínica y Salud*, 7, 19-32.
- McDonald, R. P. (1999). *Test Theory: A Unified Treatment*. Lawrence Erlbaum.
- Moral de la Rubia, J. (2019). Revisión de los criterios para validez convergente estimada a través de la Varianza Media Extraída. *Psychología*, 13, 25-41. <https://doi.org/10.21500/19002386.4119>
- Moral-de la Rubia, J. (2008). Propiedades psicométricas de la Escala de Alexitimia de Toronto de 20 reactivos en México. *Revista Electrónica de Psicología Clínica de Iztacala*, 11, 97-114. <https://www.revistas.unam.mx/index.php/rep/issue/view/151>
- Moral-de la Rubia, J. (2009). Factor Structure and Reliability of TAS-20 in Mexican Samples. *The International Journal of Hispanic Psychology*, 2, 163-176. <https://www.researchgate.net/publication/399829927>
- Moral-de la Rubia, J. (2026a). Factor Structure of the BVAQ in Mexican Married Adults. *Open Journal of Statistics*, 16, 299-327. <https://doi.org/10.4236/ojs.2026.164014>
- Moral-de la Rubia, J. (2026b). Stability and Factorial Invariance of the BVAQ-20 in Mexican Married Adults. *International Journal of Applied Research*, 12, 154-164. <https://doi.org/10.22271/allresearch.2026.v12.i8b.13996>
- Moral-de la Rubia, J., García, C. H., & Antona, C. J. (2012). Traducción y validación del Inventario Balanceado de Deseabilidad Social al Responder en una muestra probabilística de estudiantes universitarios mexicanos. *Revista de Psicología GEPU*, 3, 20-32.
- Park, C. G. (2023). Implementing Alternative Estimation Methods to Test the Construct Validity of Likert-Scale Instruments. *Korean Journal of Women Health Nursing*, 29, 85-90. <https://doi.org/10.4069/kjwhn.2023.06.14.2>

- Paulhus, D. L. (1998). *Manual for Balanced Inventory of Desirable Responding: Version 7* (BIDR-7). Multi-Health Systems.
- Paulhus, D. L. (2002). Socially Desirable Responding: The Evolution of a Construct. In H. Brau, D. Jackson, & D. E. Wiley (Eds.), *The Role of Constructs in Psychological and Educational Measurement* (pp. 46-69). Lawrence Erlbaum.
- Preece, D. A., Petrova, K., Mehta, A., Sikka, P., & Gross, J. J. (2024). Alexithymia or General Psychological Distress? Discriminant Validity of the Toronto Alexithymia Scale and the Perth Alexithymia Questionnaire. *Journal of Affective Disorders*, 352, 140-145. <https://doi.org/10.1016/j.jad.2024.01.271>
- Preece, D., Becerra, R., Robinson, K., Dandy, J., & 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>
- R Core Team (2025). *R: A Language and Environment for Statistical Computing* [Computer Software]. R Foundation for Statistical Computing. <https://www.R-project.org>
- Rousseelet, G., Pernet, C. R., & Wilcox, R. R. (2023). An Introduction to the Bootstrap: A Versatile Method to Make Inferences by Using Data-Driven Simulations. *Meta-Psychology*, 7, 1-24. <https://doi.org/10.15626/mp.2019.2058>
- Royston, J. P. (1983). Some Techniques for Assessing Multivariate Normality Based on the Shapiro-Wilk W. *Applied Statistics*, 32, 121-133. <https://doi.org/10.2307/2347291>
- Royston, P. (1992). Approximating the Shapiro-Wilk W-Test for Non-Normality. *Statistics and Computing*, 2, 117-119. <https://doi.org/10.1007/bf01891203>
- 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://bdtd.ibict.br/vufind/Record/PUC_RIO-1_47833634a519f897c8d32e71229ef89f
- Sathyanarayana, S., & Mohanasundaram, T. (2024). Fit Indices in Structural Equation Modeling and Confirmatory Factor Analysis: Reporting Guidelines. *Asian Journal of Economics, Business and Accounting*, 24, 561-577. <https://doi.org/10.9734/ajebe/2024/v24i71430>
- Schroeders, U., Kubera, F., & Gnambs, T. (2022). The Structure of the Toronto Alexithymia Scale (TAS-20): A Meta-Analytic Confirmatory Factor Analysis. *Assessment*, 29, 1806-1823. <https://doi.org/10.1177/10731911211033894>
- Sifneos, P. E. (1973). The Prevalence of 'Alexithymic' Characteristics in Psychosomatic Patients. *Psychotherapy and Psychosomatics*, 45, 255-262. <https://doi.org/10.1159/000286529>
- Sifneos, P. E. (1986). The Schalling-Sifneos Personality Scale Revised. *Psychotherapy and Psychosomatics*, 45, 161-165. <https://doi.org/10.1159/000287942>
- Sociedad Mexicana de Psicología (2010). *Código ético del psicólogo* (5a ed.). Editorial Trillas.
- Strate, L. (2025). *Assessing Alexithymia in College Students: Exploring the Psychometric Properties of the TAS-20, PAQ, and BVAQ*. Doctoral Dissertation, Tennessee Technological University. ProQuest Dissertations & Theses Global. (No. 31938608).
- Taylor, G. J., Ryan, D. P., & Bagby, R. M. (1985). Toward the Development of a New Self-Report Alexithymia Scale. *Psychotherapy and Psychosomatics*, 44, 181-199.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.
- Vorst, H. C. M., & Bermond, B. (2001). Validity and Reliability of the Bermond-Vorst Alex-

- ithymia Questionnaire. *Personality and Individual Differences*, 30, 413-434.
[https://doi.org/10.1016/s0191-8869\(00\)00033-7](https://doi.org/10.1016/s0191-8869(00)00033-7)
- Wald, A., & 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>
- 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, Article 183449092199142. <https://doi.org/10.1177/1834490921991429>
- Zech, E., Luminet, O., Rimé, B., & 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](https://doi.org/10.1002/(sici)1099-0984(199911/12)13:6<511::aid-per347>3.0.co;2-0)