

Article

Development and Preliminary Validation of a Brief Depression Scale in a Mexican Sample

Cirilo H. García-Cadena , Oscar Manuel Lara-Pinales , Ana Karen Gutiérrez-Martínez * ,
Francisco Javier Barreto Trujillo and Iván V. Martínez Alanís

Faculty of Psychology, Autonomous University of Nuevo León, Dr. Carlos Canseco 110, Mitras Centro, Monterrey 64460, Mexico; cirilo.garciacd@uanl.edu.mx (C.H.G.-C.); olarap@uanl.edu.mx (O.M.L.-P.); francisco.barretotr@uanl.edu.mx (F.J.B.T.); vladimir.martinezalns@uanl.edu.mx (I.V.M.A.)

* Correspondence: ana.gutierrezmrt@uanl.edu.mx

Abstract

Most instruments used to assess depression have been developed in sociocultural contexts that differ substantially from those of Spanish-speaking Latin American populations. To address this limitation, the present study developed and evaluated a new brief measure of depressive symptomatology for use in northeastern Mexico. The study included 414 participants ($M = 27.37$ years, $SD = 13.72$). An initial pool of 23 items, grounded in interbehavioral psychology, was developed and subsequently evaluated by a panel of 58 expert judges. Based on content-validity analyses, the instrument was reduced to 15 items. Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) supported a final four-item, one-factor solution. The final model demonstrated adequate fit to the data ($\chi^2/df = 0.281$, AGFI = 0.993, NNFI = 1.015, CFI = 1.000, RMSEA = 0.000, 90% CI [0.000, 0.094], and SRMR = 0.008). Evidence of discriminant and convergent validity was supported by a negative correlation with the Subjective Happiness Scale ($r = -0.294$, $p < 0.01$) and a positive correlation with the seven-item version of the Center for Epidemiologic Studies Depression Scale (CESD-7; $r = 0.273$, $p < 0.01$). Construct reliability was satisfactory ($H = 0.841$). Overall, the findings provide preliminary evidence supporting the reliability and validity of the proposed instrument as a brief screening measure of depressive symptomatology in a Mexican sample, mostly from Nuevo León.

Keywords: depression; depression screening; psychometric validation; interbehavioral psychology; Mexican population; scale development



Academic Editor: Domenico De Berardis

Received: 16 March 2026

Revised: 10 July 2026

Accepted: 24 July 2026

Published: 4 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Depression is a complex public health problem of global significance, affecting millions of people worldwide. According to the Institute for Health Metrics and Evaluation, approximately 280 million people are living with depression worldwide [1]. The World Health Organization reports that depression affects 5.0% of adults globally and 5.7% of adults aged 60 years and older [2]. This psychopathological disorder is more prevalent in women than in men, with estimated prevalence rates of 6% and 4%, respectively [3].

The most serious consequences of depression include suicide attempts and complete suicide. Approximately 700,000 people die by suicide each year, and suicide is currently the fourth leading cause of death among adolescents and young adults aged 15 to 29 years [2]. It has been projected that by 2030 unipolar depression will rank among the three leading contributors to the global burden of disease in high-income countries [4].

In Mexico, data from the National Survey of Self-Reported Well-being (ENBIARE), conducted by the Instituto Nacional de Estadística y Geografía (INEGI), indicate that 15.4% of the adult population reports experiencing depressive symptoms, with prevalence increasing to 19.5% among women [5]. At the state level, the highest prevalence rates have been reported in Guerrero (20.7%), Tabasco (20.6%), and Durango (19.9%), whereas the lowest rates have been observed in Quintana Roo (11.5%), Sonora (11.7%), and Nuevo León (12.3%).

1.1. Conceptual Background and Rationale for Scale Development

Given the complexity of depression and the absence of consensus regarding its etiopathogenesis, it is useful to consider several of the principal theoretical perspectives that have informed its conceptualization. The following overview is not intended to be exhaustive, but rather to summarize some of the most influential psychological approaches.

Cognitive Theory of Depression. The cognitive model of depression, originally proposed by Beck [6,7] and further elaborated by Beck and Haigh [8], posits that depression is characterized by a cognitive triad consisting of negative views of the self, the world, and the future. According to this model, the core emotional experiences associated with depression are loss and failure [8]. This perspective emphasizes internal cognitive mediators as the principal explanatory mechanisms underlying depressive symptomatology.

Psychoanalytic Conceptualization of Depression. From a psychoanalytic perspective, Sigmund Freud [9] conceptualized depression (melancholia) as a response to the loss of a loved object, accompanied by a marked decline in self-esteem and feelings of guilt. This process involves ambivalent feelings toward the lost object and the internalization of that object as part of the self. Although historically influential, this approach relies on constructs that differ substantially from those used in interbehavioral psychology.

A Radical Behaviorist Approach to Depression. Within radical behaviorism, depression is understood primarily in terms of reductions in the frequency of adaptive operant behaviors. Several authors [10–14] have emphasized that depressed individuals exhibit a marked decrease in behaviors that were previously effective and reinforced. B. F. Skinner [15] identified two principal mechanisms: (1) extinction of behaviors that were previously positively reinforced and (2) environmental changes that prevent access to reinforcement despite the preservation of the behavioral repertoire. Although this approach provides a parsimonious account, it places considerable emphasis on reinforcement processes, which represent only one component of the broader interbehavioral field.

Interbehavioral Analysis of Depression. The theoretical framework underlying the present study is based on interbehavioral psychology, primarily derived from the work of J. R. Kantor [15–19]. From this perspective, psychology examines the interbehavioral field, which consists of a set of mutually interacting factors, including stimulus objects, stimulus functions, organismic activity, response functions, the contact medium, setting factors, and interbehavioral history [17,20].

Within this framework, depression is conceptualized as a pattern of maladaptive interactions between the individual and the environment, characterized by a substantial reduction in effective and successful interactions across multiple domains, including eating, sleeping, sexual behavior, social engagement, and goal-directed activities. Behaviors such as complaining, withdrawal, and irritability become more frequent, whereas previously adaptive interactions diminish or cease. Stimulus objects that formerly functioned as effective sources of behavioral engagement no longer evoke the corresponding response functions.

Because the present study focuses on the development and validation of a new depression scale, the principal instruments currently used to assess depression are reviewed below.

The Zung Self-Rating Depression Scale. The Zung Self-Rating Depression Scale [21] is a self-administered 20-item Likert-type instrument that includes 10 positively worded and 10 negatively worded items, each rated on four response options. It assesses both psychological and somatic symptoms of depression [21]. Reported reliability estimates range from 0.73 to 0.87 across different studies [22–26].

Evidence from Spanish-speaking populations has generally supported the psychometric adequacy of the instrument, although findings have been inconsistent. For example, Cronbach's alpha coefficients have ranged from 0.548 in Colombian adolescents [26] to 0.85 in Mexican patients with type 2 diabetes [25]. These discrepancies suggest that the scale may perform better in adult populations than in children or adolescents.

A methodological concern noted by Lezama Meneses [26] is that some validation studies have used the same sample for both exploratory and confirmatory factor analyses, a practice that is not recommended because it may inflate fit indices and capitalize on chance [27,28].

Center for Epidemiologic Studies Depression Scale (CES-D). The CES-D Scale [29], its revised version (CESD-R) [30], and the abbreviated seven-item version (CESD-7) [31] are among the most widely used instruments for assessing depressive symptoms. The original CES-D consists of 20 items designed to measure depressive symptoms experienced during the previous week.

The CESD-R aligns more closely with diagnostic criteria from the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), and expands the assessment to nine symptom domains [30]. The CESD-7 retains strong psychometric properties while substantially reducing administration time. Herrero and Gracia [32] reported an internal consistency coefficient of 0.82 and a correlation of 0.92 with the full 20-item scale.

The CESD-7 has demonstrated adequate psychometric properties in diverse populations, including children, adolescents, older adults, and Mexican community samples [32–41].

Beck Depression Inventory BDI-II. The Beck Depression Inventory-II [42] is a 21-item self-report measure widely used to assess the severity of depressive symptoms. It has been extensively validated across clinical and non-clinical populations, including Spanish-speaking samples [43–49]. Due to its strong psychometric properties and broad international use, it remains one of the most widely employed instruments for depression research and clinical assessment.

An interesting research effort to create a new scale for measuring depression was carried out in Mexico by Ramos Rosales and Silva [50], with the purpose of differentiating the melancholic factor from the atypical factor. Unfortunately, they only used EFA as a validation method and inadequately employed principal component extraction, without proceeding to CFA with another sample of two hundred or more cases. The use of principal component extraction has also been heavily criticized by some authors, even going so far as to characterize it as unrelated to factor analysis, according to Tavakol and Wetzel [51].

Need for a New Depression Scale. Although several well-established depression scales are available, most were developed in sociocultural contexts that differ substantially from those of Latin American populations. Moreover, many instruments did not incorporate formal content validation by experts drawn from the same cultural context as the intended respondents. The present study sought to address this limitation by involving expert judges familiar with the Mexican sociocultural context to evaluate the relevance and clarity of the proposed items. This process was intended to ensure that the scale adequately captures culturally meaningful expressions of depressive symptomatology.

In addition, there is a practical need for a brief and psychometrically robust screening instrument that can be administered easily, including among individuals with relatively

low educational attainment. In Mexico, depression may be expressed through culturally nuanced experiences such as *desaliento* (“discouragement”), reduced motivation, and loss of interest, rather than solely through the affective experience of sadness.

Although the present scale was positively correlated with the CESD-7, the magnitude of this association was modest. This finding suggests that, while both instruments assess depressive symptomatology, they may emphasize somewhat different aspects of the construct. In particular, the present scale appears to capture the anhedonic dimension of depression more directly, as its items focus on diminished interest in activities, generalized loss of motivation, and reduced interest in life itself. These features are not explicitly represented in the CESD-7, which places greater emphasis on the affective dimension of sadness.

1.2. Main Objective

The main objective of this study was to develop a brief screening instrument capable of measuring depressive symptomatology validly and reliably in the general population of northeastern Mexico.

Specific Objectives

- 01.** *To develop an initial item pool for a depression screening scale and evaluate its content validity through expert judgment.*
- 02.** *To examine the latent factor structure of the proposed scale through EFA.*
- 03.** *To evaluate the fit and reliability of the factor structure identified through EFA using CFA in an independent subsample.*
- 04.** *To obtain preliminary evidence regarding the screening accuracy of the scale through Receiver Operating Characteristic (ROC) analysis.*

For the purposes of instrument development, the present study operationalized depression as a pattern of reduced engagement, loss of interest, diminished motivation, and negative appraisal of life context, derived from an interbehavioral framework. Based on this conceptualization, an initial pool of items was developed.

2. Materials and Methods

2.1. Participants

The study included 414 residents of the state of Nuevo Leon, Mexico. Participants ranged in age from 14 to 77 years ($M = 27.37$, $SD = 13.72$). Only 9.66% of the participants were younger than 18 years of age (i.e., between 14 and 17 years). Women comprised 67.4% of the sample ($n = 279$), and men comprised 32.6% ($n = 135$). Regarding educational attainment, 8.9% of participants had completed secondary school, 57.7% had completed high school, 22.0% held a bachelor's degree, 7.5% held a master's degree, and 3.1% held a doctoral degree. In terms of perceived socioeconomic status, 7.2% identified themselves as belonging to the high socioeconomic level, 82.4% to the middle level, and 10.4% to the low level.

2.2. Instruments

2.2.1. Sociodemographic Questionnaire

A brief sociodemographic questionnaire was administered to collect information on participants' state of residence, age, sex, educational attainment, and perceived socioeconomic status.

2.2.2. Center for Epidemiologic Studies Depression Scale (Seven-Item Version, CESD-7)

The abbreviated seven-item version of the CES-D Scale was used to obtain evidence of convergent validity for the newly developed depression scale. The Spanish version validated by Herrero and Gracia [31] assesses depressive symptoms experienced during the previous week. It is a Likert-type instrument consisting of seven items, with response options ranging from 0 to 3. Herrero and Gracia [31] reported an internal consistency coefficient (Cronbach's α) of 0.82.

In the present study, CFA supported a reduced four-item version comprising the following items: (1) "I felt that I could not shake off the sadness, even with the help of family and friends" ($\lambda = 0.82$), (2) "I felt depressed" ($\lambda = 0.85$), (3) "I felt that everything I did was an effort" ($\lambda = 0.53$), and (4) "I felt sad" ($\lambda = 0.84$). Internal consistency was adequate ($\alpha = 0.835$), and construct reliability was high ($H = 0.881$). Model fit indices were excellent: $\chi^2/df = 0.103$, GFI = 1.000, AGFI = 0.999, TLI = 1.007, CFI = 1.000, RMSEA = 0.000 (90% CI [0.000, 0.041]), and SRMR = 0.003.

2.2.3. Subjective Happiness Scale

The Subjective Happiness Scale, in its Spanish adaptation by Extremera and Fernández-Berrocal [52], was used to obtain evidence of discriminant validity for the new depression scale. This instrument assesses individuals' global subjective evaluation of their happiness. It consists of four Likert-type items with response options ranging from 1 to 7. Previous studies have reported internal consistency coefficients ranging from 0.79 to 0.94 [52]. In the present study, all four items were retained. CFA supported the proposed one-factor structure, with factor loadings ranging from 0.57 to 0.91. Internal consistency was adequate ($\alpha = 0.832$), and construct reliability was high ($H = 0.892$). Model fit indices were as follows: $\chi^2/df = 0.564$, GFI = 0.999, AGFI = 0.993, TLI = 1.004, CFI = 1.000, RMSEA = 0.000 (90% CI [0.000, 0.082]), and SRMR = 0.007.

Because both the CESD-7 and the Subjective Happiness Scale are theoretically well-established instruments with extensive psychometric support, only CFA was conducted in the present study to verify their factorial structure in the current sample [28,53].

2.3. Procedure

The development of the new depression scale followed most of the steps recommended by DeVellis [54] for scale construction and validation. First, interbehavioral psychology [18] was adopted as the theoretical framework for conceptualizing depression. Within this framework, depression is understood as a complex pattern of interactions between the individual and the environment, characterized by a marked reduction in effective and adaptive functioning across multiple areas of life.

It is conceived as an affective maladjustment, in which interactions are unsuccessful (e.g., complaining, withdrawing from others, and/or becoming angry). Important adaptive interactions diminish or disappear. Effective response functions in the person's daily and "normal" life lose their connection with stimulus functions belonging to formerly common and ordinary stimulus objects.

Based on this conceptualization, two of the authors generated an initial pool of 23 positively worded items intended to assess depressive symptomatology. At this stage, some conceptually related items were intentionally included as alternative formulations of the same theoretical indicators. This strategy was adopted to facilitate subsequent selection and refinement through expert judgment and psychometric analyses rather than to promote redundancy in the final instrument. The relevance of each item was then evaluated by a panel of 58 expert judges. Of these experts, 41.37% ($n = 24$) had completed or were pursuing doctoral studies, 18.97% ($n = 11$) had completed or were pursuing master's-level

studies, and 89.66% held a degree in psychology. Their average professional experience was 18 years, with a range from 0 to 45 years.

Content validity was assessed using Lawshe's [55] content validity ratio (CVR), which adjusts for the degree of agreement expected by chance among expert judges. Following this procedure, eight items were eliminated, resulting in a final set of 15 items for psychometric evaluation.

See Table 1, which contains the 23 items, with the eliminated items highlighted in gray, the percentage of judges who approved each item, and the content validity ratio value obtained by each item.

Table 1. Evaluation of 58 experts about the relevance of the items to measure depression.

Item	% PEA	CVR
1. I feel depressed.	70.69%	0.41
2. I don't feel like doing any activities.	79.31%	0.59
3. Life has no meaning.	91.38%	0.83
4. Life is boring.	62.07%	0.24
5. Nothing motivates me.	94.83%	0.90
6. Nothing interests me.	84.48%	0.70
7. I would like to die.	86.21%	0.72
8. The future will be horrible.	60.34%	0.21
9. I am dejected.	64.14%	0.48
10. Everything is wrong.	48.28%	0.03
11. Everything is lost.	77.59%	0.55
12. Life is horrible.	63.79%	0.28
13. Very few things in this life are truly worthwhile.	77.59%	0.55
14. Few things interest me.	74.14%	0.48
15. Sometimes I think it would be better to die.	86.21%	0.72
16. I no longer enjoy my favorite food as much as I used to.	65.86%	0.52
17. I prefer staying home that going out.	70.69%	0.41
18. It's difficult to trust others.	43.10%	0.14
19. I find gatherings with people my own age boring.	25.86%	0.48
20. Although it would be good to excel, I feel I will never be able to achieve it.	48.28%	0.03
21. I wish I could always do nothing.	62.07%	0.24
22. I have no interest in my future.	79.31%	0.59
23. I don't care whether I live or die.	81.03%	0.62

Note. Items highlighted in gray were eliminated following the content validity analysis. % PEA = Percentage of evaluators who approved the item; CVR = Content validity ratio. According to Lawshe [56], when at least 40 evaluators are used, items with a CVR of 0.29 or higher are considered valid.

Participants were recruited using a snowball sampling strategy. The authors distributed the survey through social media platforms, including WhatsApp, Instagram, and Facebook, and one author also administered the questionnaires in person. To reduce response bias, items from the various instruments were presented in randomized order within a single multiscale questionnaire.

Participation was voluntary, and all responses were collected anonymously and used exclusively for academic purposes. The study was conducted in accordance with generally accepted ethical principles for research involving human participants, including respect for autonomy, voluntary participation, confidentiality, and the minimization of potential risks. Participants were informed that their participation was voluntary and that they could discontinue the survey at any time without penalty.

Informed consent was obtained electronically from all adult participants prior to participation. Participants under 18 years of age provided informed assent before completing the survey. Because the study met the institutional criteria for risk-free research involving

an anonymous minimal-risk survey, parental or guardian consent was not required under the applicable institutional guidelines. According to Article 35, Section 1, of the Guidelines for Researchers of the Ethics Committee of the Faculty of Psychology of the Autonomous University of Nuevo León, studies meeting these criteria do not require formal ethical review and approval. Previous research has shown that participation in anonymous surveys addressing sensitive topics may involve minimal risk and can contribute valuable information regarding adolescent well-being and mental health [56–58].

2.4. Data Analysis

Data were analyzed using IBM SPSS Statistics version 25 and IBM SPSS Amos version 24. Exploratory factor analysis (EFA) was conducted in SPSS using a Pearson correlation matrix, maximum likelihood extraction, and oblimin rotation. Items were retained if they exhibited factor loadings of 0.40 or greater. The number of factors was evaluated using both the Kaiser criterion (eigenvalues greater than 1.0) and Cattell's scree test, with final decisions guided by the interpretability of the resulting factor structure. Pearson product-moment correlations were computed to evaluate discriminant validity (correlations between the new depression scale and the Subjective Happiness Scale) and convergent validity (correlations between the new depression scale and the CESD-7).

Construct reliability (H) was calculated using the spreadsheet provided by Domínguez-Lara [59]. Cronbach's alpha (α) was computed in SPSS, and McDonald's omega (Ω) was also estimated as an additional indicator of internal consistency. Receiver Operating Characteristic (ROC) analysis was conducted to estimate the area under the curve (AUC), determine the optimal cutoff score, and calculate sensitivity, specificity, and the Youden index ($J = \text{sensitivity} + \text{specificity} - 1$). Finally, average variance extracted (AVE) was calculated as an indicator of internal convergent validity.

3. Results

3.1. Validation of the New Depression Scale Through EFA

The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was 0.86, indicating that the data were suitable for EFA. In addition, the substantial inter-item correlations suggested the presence of underlying latent dimensions [60]. Bartlett's test of sphericity was statistically significant ($\chi^2 = 1133.327$, $p < 0.001$), indicating that the correlation matrix differed significantly from an identity matrix and was therefore appropriate for factor extraction. The number of factors were evaluated using both the Kaiser criterion and Cattell's scree test [61]. As shown in Figure 1, the scree plot exhibited a clear inflection after the third eigenvalue, suggesting the retention of three factors.

Although three factors initially emerged, items exhibiting substantial cross-loadings were removed following the recommendations of Costello and Osborne [62] and Matsunaga [63]. Specifically, items were excluded when they showed secondary loadings ≥ 0.30 or when the difference between primary and secondary loadings was less than 0.20. This procedure was implemented to ensure a simple and theoretically interpretable factor structure.

Of the 15 items that remained after content validation, only four met the retention criteria established for the EFA. The retained items and their corresponding factor loadings were: i2 ("I don't feel like doing activities"; $\lambda = 0.498$), i3 ("Life has no meaning"; $\lambda = 0.620$), i4 ("Nothing motivates me"; $\lambda = 0.770$), and i5 ("Nothing interests me"; $\lambda = 0.831$).

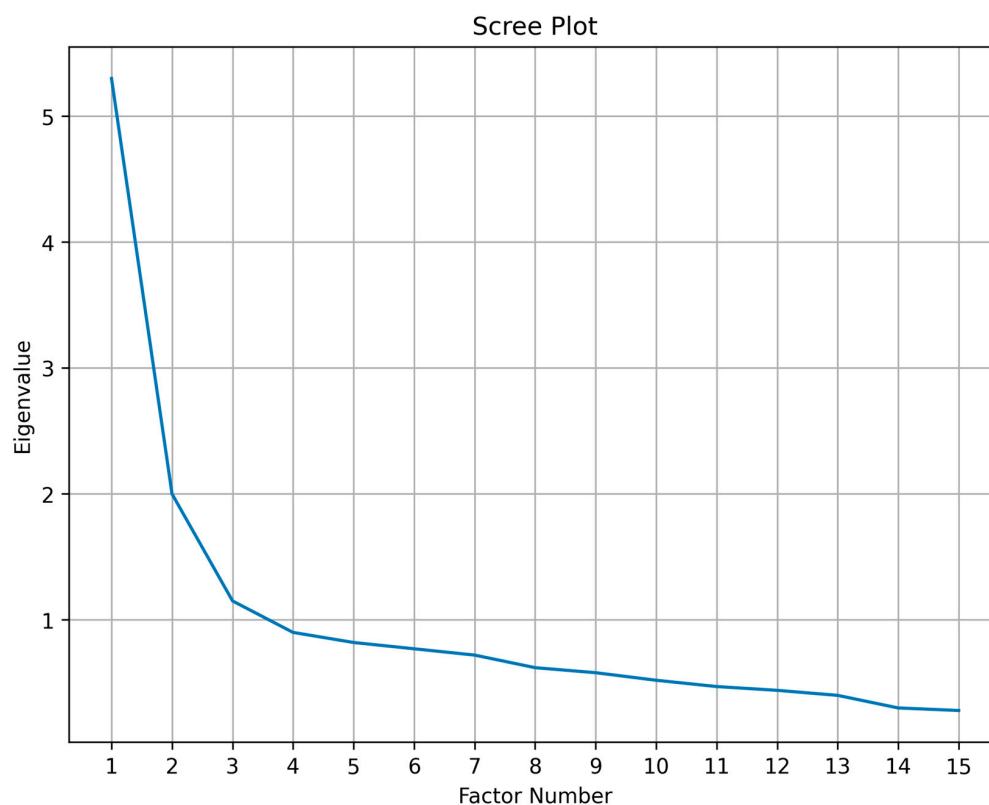


Figure 1. Factor structure of the new depression scale derived from EFA.

3.2. Validation of the New Depression Scale Through CFA

The factor structure identified through EFA was subsequently evaluated using CFA in an independent subsample of 207 participants. All four items retained in the EFA were preserved because they demonstrated satisfactory factor loadings and theoretical coherence with the construct of depression. Table 2 presents the final measurement model, including descriptive statistics and standardized factor loadings.

Table 2. Validated factor structure of the new depression scale.

Items	<i>M</i>	<i>SD</i>	<i>K</i>	<i>SK.</i>	λ
i2—I don't feel like doing activities.	2.29	1.587	−1.213	0.117	0.64
i3—Life has no meaning.	2.18	1.836	−1.440	0.231	0.71
i4—Nothing motivates me.	2.09	1.676	−1.148	0.384	0.82
i5—Nothing interests me.	1.68	1.608	−0.663	0.751	0.78
Multivariate normality	5.898				

Note. Validation was performed using CFA with maximum likelihood extraction method ($n = 207$). λ = Factor loading; *M* = Mean; *SD* = Standard deviation; *K* = Kurtosis; *SK.* = Skewness. The scale's goodness-of-fit indices were $\chi^2/df = 0.281$, GFI = 0.999, AGFI = 0.993, NNFI = 1.015, CFI = 1.000, RMSEA = 0.000 (90% CI, 0.000–0.094) and SRMR = 0.008.

The standardized factor loadings ranged from 0.64 to 0.82, indicating that all items made substantial contributions to the latent construct. No evidence of problematic skewness or kurtosis was observed. The multivariate normality coefficient was 5.898, which was considered acceptable for estimation using the maximum likelihood method.

The one-factor model demonstrated excellent fit to the data: $\chi^2/df = 0.281$, GFI = 0.999, AGFI = 0.993, NNFI = 1.015, CFI = 1.000, RMSEA = 0.000 (90% CI [0.000, 0.094]), and SRMR = 0.008. Collectively, these indices provide strong support for the adequacy of the proposed measurement model.

3.3. Internal Consistency of the New Depression Scale

The internal consistency of the final four-item depression scale was satisfactory across all reliability indicators examined. Construct reliability was high ($H = 0.841$), while McDonald's omega ($\Omega = 0.828$) and Cronbach's alpha ($\alpha = 0.826$) also indicated adequate reliability. Furthermore, all standardized factor loadings exceeded the recommended threshold of 0.50, ranging from 0.64 to 0.82, providing additional evidence of internal convergent validity [64].

3.4. Convergent and Discriminant Validity

Evidence of discriminant validity was obtained through the negative association between scores on the new depression scale and scores on the Subjective Happiness Scale [65]. As expected, individuals reporting higher levels of depressive symptomatology tended to report lower levels of subjective happiness ($r = -0.294$, $p < 0.01$). Although modest in magnitude, this correlation reached the minimum effect size recommended by Ferguson [66].

Evidence of convergent validity was also observed. Scores on the new depression scale correlated positively and significantly with scores on the CESD-7 ($r = 0.273$, $p < 0.01$), indicating that both instruments assess related aspects of depressive symptomatology. Although the magnitude of the correlation was relatively small, it was consistent with theoretical expectations and met the minimum effect-size criterion proposed by Ferguson [66].

Additional support for convergent validity was provided by the average variance extracted ($AVE = 0.55$), which exceeded the recommended minimum value of 0.50 [65]. Likewise, the magnitude of the factor loadings further supported the conclusion that the items adequately represented the underlying construct [67].

3.5. Receiver Operating Characteristic (ROC) Analysis

The ROC analysis yielded an area under the curve (AUC) of 0.69, indicating modest discriminatory accuracy for screening purposes (Figure 2).

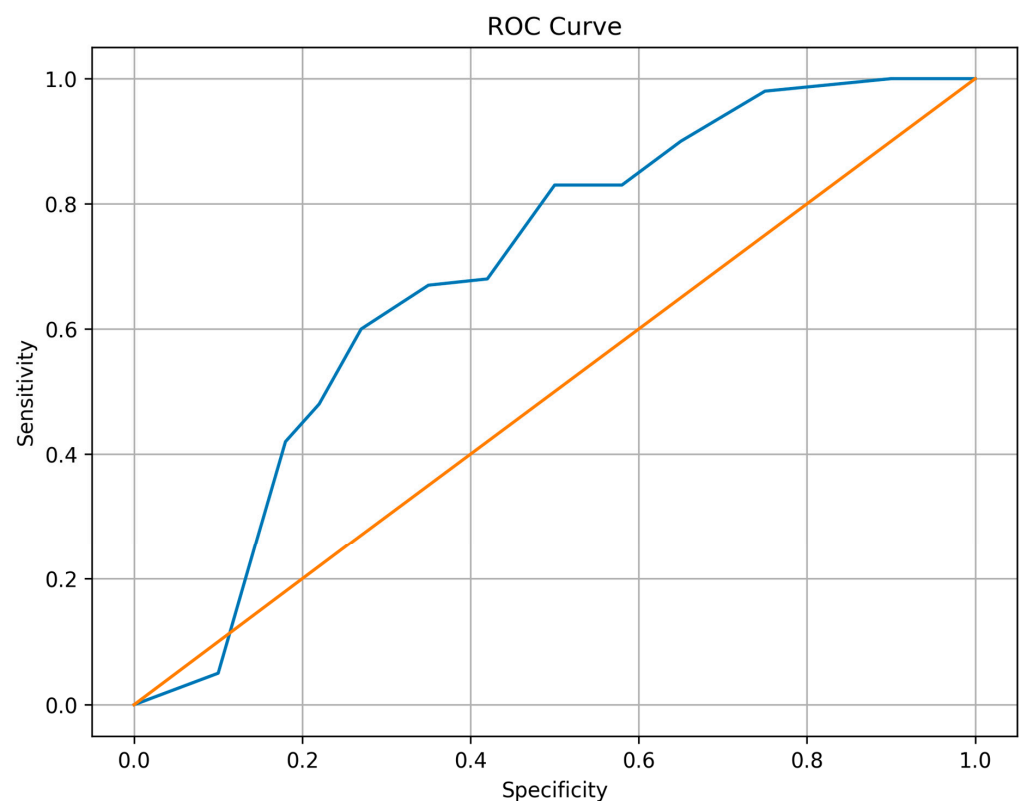


Figure 2. Receiver operating characteristic (ROC) curve. The blue line represents the ROC curve, and the orange diagonal line represents the reference line corresponding to chance-level classification.

The optimal cutoff score was identified using the Youden index. The highest value of the index ($J = 0.20$) was obtained at a cutoff score of 8.5, corresponding to a sensitivity of 0.75 and a specificity of 0.45. This cutoff point provided the most favorable balance between sensitivity and specificity among the evaluated cutoff scores, although specificity remained relatively low.

Given that the instrument was developed as a screening tool rather than a diagnostic instrument, greater emphasis was placed on sensitivity than specificity. Consequently, the cutoff score of 8.5 was considered the most appropriate for identifying individuals who may require further psychological or psychiatric evaluation, rather than for diagnostic classification. The comparative results for the different cutoff scores are presented in Table 3.

Table 3. Comparison of area under the curve, confidence interval, Youden index, sensitivity, specificity and cut-off points in the new Mexican depression scale.

CP	AUC	95%CI	J	Se.	Sp.
7.5	0.65	[0.57–0.73]	0.19	0.69	0.50
8.5	0.69	[0.57–0.81]	0.20	0.75	0.45
9.5	0.67	[0.50–0.93]	0.00	0.60	0.40

Note. CP = Cutoff point; AUC = Area under the curve; CI = Confidence interval; J = Youden index; Se. = Sensitivity; Sp. = Specificity.

4. Discussion

4.1. Main Psychometric Findings

The present study provides converging evidence supporting the validity and reliability of the proposed instrument as a brief measure of depressive symptomatology for use in populations from northeastern Mexico. Multiple sources of evidence and complementary psychometric procedures were employed throughout the scale-development process.

Initially, an item pool of 23 statements was generated based on the theoretical framework of interbehavioral psychology. Subsequently, expert judgment was used to evaluate the representativeness and cultural relevance of the items, thereby strengthening content validity [68]. The resulting item set was then subjected to exploratory and confirmatory factor analyses [28,54], which supported a unidimensional (single-factor) latent structure. Therefore, the final instrument should be interpreted as a unidimensional measure of depressive symptomatology.

Additional evidence of construct validity was obtained through analyses of convergent and discriminant validity. As expected, scores on the new scale were positively associated with scores on the CESD-7, a theoretically related measure of depressive symptomatology, and negatively associated with scores on the Subjective Happiness Scale (SHS) [65]. According to Campbell and Fiske [69], both forms of validity provide important evidence regarding construct interpretation. However, the magnitude of the observed correlations was relatively modest, suggesting that the findings provide limited but statistically significant preliminary support for convergent and discriminant validity. Furthermore, the average variance extracted (AVE) exceeded recommended thresholds, indicating that the retained items adequately represented the underlying latent construct [64,67].

Receiver operating characteristic (ROC) analyses provided preliminary evidence regarding the instrument's capacity to distinguish individuals with elevated depressive symptomatology from those with lower symptom levels. Although the scale was designed primarily as a screening instrument, the results suggest a modest level of discriminatory accuracy that may be useful for preliminary screening purposes. However, these findings should be interpreted cautiously because the ROC analyses were based on a self-report external criterion rather than an independent clinical diagnosis.

Finally, evidence of score reliability was consistently supported across multiple indicators. Internal consistency coefficients (Cronbach's α and McDonald's Ω) were satisfactory, and construct reliability (H) exceeded recommended values. Taken together, these findings suggest that the proposed instrument provides a reliable assessment of depressive symptomatology and possesses an adequate psychometric foundation for screening and research purposes. Higher scores on the scale indicate greater depressive symptomatology, whereas lower scores indicate lower levels of depressive symptomatology.

4.2. Theoretical Significance

The development of new instruments to assess psychological and psychopathological constructs remains an important scientific endeavor, even in areas where numerous measures are already available. This is particularly true in the case of depression, one of the most extensively studied constructs in clinical psychology, health psychology, and psychiatry.

One reason for continuing to develop new instruments is that many existing scales were created under methodological standards that differ substantially from current recommendations for instrument development and validation. As noted by Hinkin [70], a considerable number of psychological measures were developed using procedures that would now be regarded as limited or insufficiently rigorous. Contemporary psychometric standards emphasize the importance of clearly defining the construct domain, generating theoretically grounded items, obtaining evidence of content validity, and conducting comprehensive evaluations of the internal structure of the instrument [28,54].

From a theoretical perspective, the present study contributes to literature by developing a depression measure derived explicitly from interbehavioral psychology. To our knowledge, no previously published depression scale has been constructed from this conceptual framework. Consequently, the proposed instrument contributes to expanding the empirical applications of interbehavioral theory while providing an alternative perspective for understanding and assessing depressive symptomatology.

The content of the final four items can also be interpreted from an interbehavioral perspective. Items referring to loss of interest, motivation, and willingness to engage in activities reflect a disruption in the functional relations that historically connect environmental events with effective patterns of responding. In this sense, activities, objects, and social situations that previously acquired reinforcing or functional significance lose their capacity to regulate behavior. Likewise, the item referring to a perceived lack of meaning in life may represent a more severe deterioration of these functional relations, reflecting a substantial reduction in the environmental functions that support adaptive engagement with everyday life.

A second consideration concerns the cultural context in which psychological assessment takes place. In Mexico, as in many Spanish-speaking countries, depression is commonly assessed using instruments originally developed in sociocultural contexts that differ substantially from those in which they are ultimately applied. Although translation and adaptation procedures can improve cross-cultural applicability, the meaning and interpretation of questionnaire items are inevitably influenced by cultural, linguistic, and contextual factors.

Accordingly, one potential advantage of the present instrument is that it was developed and evaluated within the same sociocultural context in which it is intended to be used. The incorporation of expert judges familiar with the target population strengthened content validity and helped ensure that the retained items reflected culturally meaningful expressions of depressive symptomatology [27]. Consequently, the scale may offer a culturally relevant alternative for the assessment of depression in populations from northeastern Mexico, particularly in screening and research contexts.

More broadly, the present findings support the value of integrating theoretical coherence, cultural relevance, and contemporary psychometric procedures during instrument development. Such an approach may contribute to the creation of measures that are both scientifically robust and contextually appropriate for the populations in which they are applied.

4.3. Methodological Limitations and Practical Implications

Although the present findings provide support for the psychometric adequacy of the instrument, several methodological considerations should be considered when interpreting the results.

First, the cross-sectional design precludes conclusions regarding temporal stability beyond internal consistency estimates and does not permit causal inferences regarding the construct under investigation. Future longitudinal studies would allow for the examination of predictive validity, temporal stability, and sensitivity to change.

Second, although the factorial analyses supported the proposed unidimensional structure, alternative models (e.g., bifactor or higher-order structures) could be explored to determine whether the construct is better represented by a dominant general factor or by multiple substantively meaningful dimensions. Likewise, testing measurement invariance across sex, age groups, and other relevant sociodemographic characteristics would provide additional evidence regarding the structural robustness of the instrument.

Third, the exclusive reliance on self-report measures introduces the possibility of shared-method variance and response bias. Future studies incorporating multimethod assessment strategies, including clinician ratings and behavioral indicators, could provide stronger evidence of construct validity and reduce the influence of mono-method bias.

Fourth, although the ROC analyses indicated acceptable discriminatory performance, replication in clinical samples with formally established diagnoses is needed to further evaluate diagnostic utility and refine cutoff scores.

From a theoretical perspective, depressive symptomatology may be more appropriately understood as a dynamic pattern of interactions between individuals and their environments rather than as a static entity. Future research grounded in this perspective may help clarify whether the observed factor structure reflects latent traits, functional response classes, or context-dependent interactional patterns.

It should also be noted that, given the small number of indicators and the relative simplicity of the final measurement model, the excellent CFA fit indices should be interpreted cautiously. Although these results provide support for the adequacy of the proposed structure, they should not be regarded as sufficient evidence of broad construct validity or comprehensive coverage of depressive symptomatology. Taken together, the present findings provide preliminary evidence of reliability and structural validity. However, stronger conclusions regarding the generalizability and practical utility of the instrument will require replication across diverse populations, tests of measurement invariance, and corroboration through multimethod assessment approaches.

Several additional limitations should also be acknowledged. The sample was obtained through non-probability sampling and consisted primarily of participants from Nuevo Leon and other areas of northeastern Mexico, which limits the generalizability of the findings to other Mexican populations. Furthermore, women were overrepresented in the sample, and individuals with lower educational attainment were relatively underrepresented, both of which may have influenced the psychometric properties observed.

Another limitation concerns the absence of test-retest analyses, which prevented evaluation of the temporal stability of the instrument. In addition, although the final four-item version demonstrated satisfactory psychometric performance, its brevity may limit the

extent to which it captures the multidimensional complexity of depressive symptomatology. Further research on factorial invariance is warranted to determine whether the scale measures the same construct equivalently across different groups and time points, thereby supporting meaningful comparisons of latent means, observed scores, and factor loadings [71].

An additional limitation relates to the relatively low specificity observed in the ROC analyses (0.45). In addition, the ROC analyses were conducted using another self-report measure as the external criterion rather than an independent clinical diagnosis, which limits the interpretation of the obtained classification indices and should be considered when evaluating the reported screening performance. Nevertheless, this result should be interpreted considering the intended purpose of the instrument. Because the scale was designed as a screening rather than a diagnostic measure, prioritizing sensitivity may be justified to reduce the likelihood of failing to identify individuals with clinically relevant depressive symptomatology. In this context, the observed sensitivity (0.75) suggests a reasonable capacity to detect probable cases, albeit at the cost of an increased false-positive rate. Consequently, positive screening results should be interpreted as preliminary indicators requiring further clinical assessment rather than as evidence of a definitive diagnosis.

Future studies should also evaluate criterion validity using external indicators of depression, including structured clinical interviews and professional diagnoses [72]. Finally, the application of item response theory (IRT) models [73] could provide more detailed information regarding item functioning across different levels of depressive severity, thereby complementing the classical test theory approach adopted in the present study.

Author Contributions: Conceptualization, C.H.G.-C. and I.V.M.A.; methodology, C.H.G.-C.; software, O.M.L.-P.; validation, C.H.G.-C., O.M.L.-P. and A.K.G.-M.; formal analysis, C.H.G.-C. and O.M.L.-P.; investigation, O.M.L.-P., F.J.B.T. and I.V.M.A.; resources, F.J.B.T.; data curation, O.M.L.-P.; writing—original draft preparation, C.H.G.-C.; writing—review and editing, A.K.G.-M.; visualization, A.K.G.-M.; supervision, C.H.G.-C.; project administration, A.K.G.-M.; funding acquisition, not applicable. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: According to the Guidelines for Researchers of the Ethics Committee of the Faculty of Psychology, Autonomous University of Nuevo León, specifically Article 35, Section 1, the present study met the criteria for risk-free research which consisted of an anonymous, voluntary, minimal-risk survey that did not collect identifiable personal information. Consequently, formal ethical review and approval were not required.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request. The data are not publicly available to protect participant anonymity.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT-5.5 (OpenAI) for translation, language editing, and assistance with reference formatting. All AI-generated outputs were carefully reviewed and edited by the authors, who assume full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AUC	Area under the curve
AVE	Average variance extracted
CFA	Confirmatory factor analysis

CES-D	Center for Epidemiologic Studies Depression Scale
CESD-7	Seven-item version of the Center for Epidemiologic Studies Depression Scale
CESD-R	Revised Center for Epidemiologic Studies Depression Scale
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
EFA	Exploratory factor analysis
INEGI	National Institute of Statistics and Geography (Mexico)
ROC	Receiver operating characteristic
SHS	Subjective Happiness Scale

References

1. Institute for Health Metrics and Evaluation. Global Health Data Exchange (GHDx). Available online: <https://vizhub.healthdata.org/gbd-results/> (accessed on 9 December 2025).
2. World Health Organization. Depression: Fact Sheet. Available online: <https://www.who.int/es/news-room/fact-sheets/detail/depression> (accessed on 8 September 2025).
3. Woody, C.A.; Ferrari, A.J.; Siskind, D.J.; Whiteford, H.A.; Harris, M.G. Prevalence and incidence of perinatal depression. *J. Affect. Disord.* **2017**, *219*, 86–92. [CrossRef] [PubMed]
4. Mathers, C.D.; Loncar, D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* **2006**, *3*, e442. [CrossRef] [PubMed]
5. Instituto Nacional de Estadística y Geografía. *Encuesta Nacional Sobre Bienestar Autorreportado (ENBIARE) 2021*; INEGI: Aguascalientes, Mexico, 2021. Available online: <https://www.inegi.org.mx/programas/enbiare/2021/> (accessed on 7 October 2025).
6. Beck, A.T. *Depression: Clinical, Experimental and Theoretical Aspects*; Harper & Row: New York, NY, USA, 1967.
7. Beck, A.T. Cognitive models of depression. *J. Cogn. Psychother.* **1987**, *1*, 5–37.
8. Beck, A.T.; Haigh, E.A.P. Advances in cognitive theory and therapy: The generic cognitive model. *Annu. Rev. Clin. Psychol.* **2014**, *10*, 1–24. [CrossRef] [PubMed]
9. Freud, S. Mourning and melancholia. In *On Metapsychology: The Theory of Psychoanalysis*; Dickson, A., Ed.; Penguin Books: London, UK, 1991; pp. 59–98.
10. Ferster, C.B. Animal behavior and mental illness. *Psychol. Rec.* **1966**, *16*, 345–356. [CrossRef]
11. Ferster, C.B. A functional analysis of depression. *Am. Psychol.* **1973**, *28*, 857–870. [CrossRef] [PubMed]
12. Lazarus, A.A. Learning theory and the treatment of depression. *Behav. Res. Ther.* **1968**, *6*, 83–89. [CrossRef] [PubMed]
13. Lewinsohn, P.M.; Weinstein, M.S.; Shaw, D.A. Depression: A clinical research approach. In *Advances in Behavior Therapy: 1968*; Rubin, R.D., Franks, C.M., Eds.; Academic Press: New York, NY, USA, 1969; pp. 231–240.
14. Skinner, B.F. The operant side of behavior therapy. *J. Behav. Ther. Exp. Psychiatry* **1988**, *19*, 171–179. [CrossRef] [PubMed]
15. Ribes, E. Kantor's contribution to psychology, or what is behavior? *Behav. Anal.* **1988**, *23*, 94–100.
16. Kantor, J.R. *Principles of Psychology*; Principia Press: Bloomington, IN, USA, 1924; Volume 1.
17. Kantor, J.R. *Interbehavioral Psychology*; Principia Press: Bloomington, IN, USA, 1958.
18. Kantor, J.R.; Smith, N. *The Science of Psychology: An Interbehavioral Survey*; Principia Press: Bloomington, IN, USA, 1975.
19. Kantor, J.R. Psychological retardation and interbehavioral maladjustments. *Psychol. Rec.* **1982**, *32*, 305–313. [CrossRef]
20. Hayes, L.J.; Fryling, M.J. *Interbehaviorism: A Comprehensive Guide to the Foundations of Kantor's Theory and Its Applications for Modern Behavior Analysis*; Context Press: Reno, NV, USA, 2023.
21. Zung, W.W.K. A self-rating depression scale. *Arch. Gen. Psychiatry* **1965**, *12*, 63–70. [CrossRef] [PubMed]
22. Jokelainen, J.; Timonen, M.; Keinänen-Kiukaanniemi, S.; Härkönen, P.; Jurvelin, H.; Suija, K. Validation of the Zung self-rating depression scale in older adults. *Scand. J. Prim. Health Care* **2019**, *37*, 353–357. [CrossRef] [PubMed]
23. Jonghe, J.F.; Baneke, J.J. The Zung self-rating depression scale: A replication study on reliability, validity and prediction. *Psychol. Rep.* **1989**, *64*, 833–834. [CrossRef]
24. Dunstan, D.A.; Scott, N. Clarification of the cut-off score for Zung's self-rating depression scale. *BMC Psychiatry* **2019**, *19*, 177. [CrossRef] [PubMed]
25. López-Pérez, M.G.; Ávila-Jiménez, L. Propiedades psicométricas de una escala de depresión en pacientes con diabetes tipo 2. *Rev. Electrón. Psicol. Iztacala* **2022**, *25*, 1481–1498.
26. Lezama Meneses, S.R. Propiedades psicométricas de la escala de Zung en adolescentes colombianos. *Psicología* **2012**, *6*, 91–101. [CrossRef]
27. Boateng, G.O.; Neilands, T.B.; Frongillo, E.A.; Melgar-Quinonez, H.R.; Young, S.L. Best practices for developing and validating scales. *Front. Public Health* **2018**, *6*, 149. [CrossRef] [PubMed]
28. Brown, T.A. *Confirmatory Factor Analysis for Applied Research*, 2nd ed.; Guilford Press: New York, NY, USA, 2015.

29. Radloff, L.S. The CES-D scale: A self-report depression scale for research in the general population. *Appl. Psychol. Meas.* **1977**, *1*, 385–401. [[CrossRef](#)]
30. Eaton, W.W.; Smith, C.; Ybarra, M.; Muntaner, C.; Tien, A. Center for Epidemiologic Studies Depression Scale: Review and revision. In *The Use of Psychological Testing for Treatment, Planning and Outcomes Assessment*, 3rd ed.; Maruish, M.E., Ed.; Lawrence Erlbaum: Mahwah, NJ, USA, 2004; Volume 3, pp. 363–377.
31. Herrero, J.; Gracia, E. Una medida breve de la sintomatología depresiva (CESD-7). *Salud Ment.* **2007**, *30*, 40–46.
32. Benjet, C.; Guzmán, L.; Quintanilla, G.; Roque, A.; León, R. Validez y confiabilidad de la escala CES-D en peripúberes. *Rev. Mex. Psicol.* **1999**, *16*, 175–185.
33. Camacho, P.A.; Rueda-Jaimes, G.E.; Latorre, J.F.; Navarro-Mancilla, A.A. Validez y confiabilidad de la escala CES-D en adolescentes colombianos. *Biomédica* **2009**, *29*, 260–269. [[CrossRef](#)]
34. Cárcamo-Zepeda, E.; Valencia, P.D. Propiedades psicométricas del CESD-7 en adolescentes mexicanos. *Rev. Psicol. Clín. Niños Adolesc.* **2025**, *12*, 196–204. [[CrossRef](#)]
35. González-Forteza, C.; Jiménez, J.A.; Ramos-Lira, L.; Wagner, F. Aplicación de la escala CES-D en adolescentes de la Ciudad de México. *Salud Pública Méx.* **2008**, *50*, 292–299. [[CrossRef](#)] [[PubMed](#)]
36. González-Forteza, C.; Solís Torres, C.; Jiménez Tapia, A.; Hernández Fernández, I.; González-González, A.; Juárez García, F.; Medina-Mora, M.E.; Fernández-Varela Mejía, H. Confiabilidad y validez de la escala CES-D. *Salud Ment.* **2011**, *34*, 1–7.
37. Merino-Soto, C.; Salas-Blas, E.; Pérez-Amezcu, B.; García-Rivas, J.; Peña, O.I.G.; Toledano-Toledano, F. Brief Sensations Seeking Scale (BSSS): Validity evidence in Mexican adolescents. *Int. J. Environ. Res. Public Health* **2022**, *19*, 7978. [[CrossRef](#)] [[PubMed](#)]
38. Salas-Blas, E.; Merino-Soto, C.; Pérez-Amezcu, B.; Toledano-Toledano, F. Social Networks Addiction (SNA-6)–Short. *Front. Psychol.* **2021**, *12*, 774847. [[CrossRef](#)] [[PubMed](#)]
39. Reyes, M.; Soto, A.; Milla, J.; García-Ramírez, A.; Hubard, L.; Mendoza, H.; Mejía, L.; García, M.C.; Wagner, F.A. Actualización de la escala CES-D en una muestra geriátrica mexicana. *Salud Ment.* **2003**, *26*, 59–68.
40. Sánchez-García, S.; García-Domínguez, A.; García-Peña, C.; Gutiérrez-Gutiérrez, L.; Narváez-Macias, L.; Juárez-Cedillo, T. Utilización de la versión reducida de la escala CES-D en adultos mayores mexicanos. *Entreciencias* **2014**, *2*, 137–150. [[CrossRef](#)]
41. Salgado, N. Características psicométricas de la escala de depresión del Centro de Estudios Epidemiológicos en mujeres mexicanas adultas en áreas rurales. *Salud Pública Méx.* **1994**, *36*, 200–209.
42. Beck, A.T.; Steer, R.A.; Brown, G.K. *BDI-II: Beck Depression Inventory—Second Edition Manual*; Psychological Corporation: San Antonio, TX, USA, 1996. [[CrossRef](#)]
43. García-Cadena, C.H.; Daniel-González, L.; Lara Pinales, O.M.; Quiceno Sierra, J.M. Propensión a corrupción y su efecto sobre la depresión: Un modelo psicológico. *Acción Psicológica* **2025**, *22*, 69–81. [[CrossRef](#)]
44. Sanz, J.; Navarro, M.E.; Vázquez, C. Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): Propiedades psicométricas en estudiantes universitarios. *Anal. Modif. Conducta* **2003**, *29*, 239–288.
45. Sanz, J.; García-Vera, M.P.; Espinosa, R.; Fortún, M.; Vázquez, C. Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): Propiedades psicométricas en pacientes con trastornos psicológicos. *Clín. Salud* **2005**, *16*, 121–142.
46. Osman, A.; Kopper, B.A.; Barrios, F.; Gutierrez, P.M.; Bagge, C.L. Reliability and validity of the Beck Depression Inventory–II with adolescent psychiatric inpatients. *Psychol. Assess.* **2004**, *16*, 120–132. [[CrossRef](#)] [[PubMed](#)]
47. Wang, Y.P.; Gorenstein, C. Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Rev. Bras. Psiquiatr.* **2013**, *35*, 416–431. [[CrossRef](#)] [[PubMed](#)]
48. Rosas-Santiago, F.J.; Rodríguez-Pérez, V.; Hernández-Aguilera, R.D.; Lagunes-Córdoba, R. Estructura factorial del Inventario de Depresión de Beck II. *Salud Uninorte* **2020**, *36*, 436–449. [[CrossRef](#)]
49. García-Batista, Z.E.; Guerra-Peña, K.; Cano-Vindel, A.; Herrera-Martínez, S.X.; Medrano, L.A. Validity and reliability of the Beck Depression Inventory-II. *PLoS ONE* **2018**, *13*, e0199750. [[CrossRef](#)] [[PubMed](#)]
50. Ramos Rosales, J.H.; Silva, C. Construcción y validación de una escala para evaluar subtipos del trastorno depresivo mayor. *Enseñ. Investig. Psicol.* **2016**, *21*, 89–95.
51. Tavakol, M.; Wetzel, A. Factor analysis: A means for theory and instrument development. *Int. J. Med. Educ.* **2020**, *11*, 245–247. [[CrossRef](#)] [[PubMed](#)]
52. Extremera, N.; Fernández-Berrocal, P. Spanish adaptation of the Subjective Happiness Scale (SHS). *Span. J. Psychol.* **2014**, *17*, E83. [[CrossRef](#)] [[PubMed](#)]
53. Izquierdo, I.; Olea, J.; Abad, F.J. Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema* **2014**, *26*, 395–400. [[CrossRef](#)] [[PubMed](#)]
54. DeVellis, R.F. *Scale Development: Theory and Applications*, 4th ed.; Sage: Thousand Oaks, CA, USA, 2016.
55. Lawshe, C.H. A quantitative approach to content validity. *Pers. Psychol.* **1975**, *28*, 563–575. [[CrossRef](#)]
56. Blades, C.A.; Stritzke, W.G.K.; Page, A.C.; Brown, J.D. The benefits and risks of asking research participants about suicide: A meta-analysis. *Clin. Psychol. Rev.* **2018**, *64*, 1–12. [[CrossRef](#)] [[PubMed](#)]

57. Neelakantan, L.; Fry, D.; Florian, L.; Meinck, F. Adolescents' experiences of participating in sensitive research. *Trauma Violence Abus.* **2023**, *24*, 1405–1426. [[CrossRef](#)] [[PubMed](#)]
58. Polihronis, C.; Cloutier, P.; Kaur, J.; Skinner, R.; Cappelli, M. What's the harm in asking? *Arch. Suicide Res.* **2022**, *26*, 325–347. [[CrossRef](#)] [[PubMed](#)]
59. Domínguez-Lara, S.A. Evaluación de la confiabilidad del constructo mediante el coeficiente H. *Psychologia* **2016**, *10*, 87–94. [[CrossRef](#)]
60. Revelle, W. *psych: Procedures for Psychological, Psychometric, and Personality Research*; R Package Version 2.5; Northwestern University: Evanston, IL, USA, 2020. Available online: <https://cran.r-project.org/package=psych> (accessed on 21 November 2025).
61. Cattell, R.B. The scree test for the number of factors. *Multivar. Behav. Res.* **1966**, *1*, 245–276. [[CrossRef](#)] [[PubMed](#)]
62. Costello, A.B.; Osborne, J. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* **2005**, *10*, 7.
63. Matsunaga, M. How to factor-analyze your data right: Do's, don'ts, and how-to's. *Int. J. Psychol. Res.* **2010**, *3*, 98–111. [[CrossRef](#)]
64. Hair, J.F.; Hult, G.T.M.; Ringle, C.M.; Sarstedt, M. *A Primer on Partial Least Squares Structural Equation Modeling*; Sage: Thousand Oaks, CA, USA, 2014.
65. Lyubomirsky, S.; Lepper, H.S. A measure of subjective happiness: Preliminary reliability and construct validation. *Soc. Indic. Res.* **1999**, *46*, 137–155. [[CrossRef](#)]
66. Ferguson, C.J. An effect size primer: A guide for clinicians and researchers. *Prof. Psychol. Res. Pract.* **2009**, *40*, 532–538. [[CrossRef](#)]
67. Wu, X.; Liu, H.; Zhang, C.; Zhang, F.; Xie, B.; Zhong, X. Using structural equation modeling to investigate students' satisfaction. *BMC Med. Educ.* **2024**, *24*, 792. [[CrossRef](#)] [[PubMed](#)]
68. Lynn, M.R. Determination and quantification of content validity. *Nurs. Res.* **1986**, *35*, 382–386. [[CrossRef](#)]
69. Campbell, D.T.; Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* **1959**, *56*, 81–105. [[CrossRef](#)]
70. Hinkin, T.R. A review of scale development practices in the study of organizations. *J. Manag.* **1995**, *21*, 967–988. [[CrossRef](#)]
71. Leitgöb, H.; Koch, A.; Davidov, E.; Schmidt, P. Measurement invariance in the social sciences. *Int. J. Intercult. Relat.* **2023**, *93*, 102805. [[CrossRef](#)] [[PubMed](#)]
72. Messick, S. Validity of psychological assessment: Validation of inferences from persons' responses and performances. *Am. Psychol.* **1995**, *50*, 741–749. [[CrossRef](#)]
73. Shaw, A.; Kapnek, M.; Morelli, N.A. Measuring creative self-efficacy: An item response theory analysis. *Front. Psychol.* **2021**, *12*, 678033. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.