ASIDE HEALTHCARE

ASIDE Health Sciences

ASIDE HEALTH SCIENCES

Review Article

Cross-Cultural Application of the Beck Depression Inventory (BDI-I/II) for Assessing Adolescent Depression in Low- and Middle-Income Countries: A Critical Review of Evidence from the Last Quarter-Century

Anas Al-Qatefy[1,*], Abdelrahman Tolba[2], Zeyad Eldawy[2], Mahmoud Ghallab[3]

1-Faculty of Medicine, Mansoura National University, Mansoura, Egypt
2-Neurology Department, Faculty of Medicine, Alexandria University, Alexandria, Egypt
3-Psychiatry Department, Faculty of Medicine, Mansoura National University, Mansoura, Egypt

ARTICLE INFO

Article history:
Received 9 Dec. 2025
Received in revised form 28 Dec. 2025
Accepted 19 Jan. 2026
Published 22 Jan. 2026

Keywords:
Beck Depression Inventory (BDI)
Adolescent Depression
Cross-Cultural Validation
Low- and Middle-Income Countries (LMICs)
Measurement Invariance
Psychiatry
Psychology

ABSTRACT

Background: The Beck Depression Inventory (BDI-I/II) versions are widely used for identifying depression among adolescents in low- and middle-income countries (LMICs). Nevertheless, their global usage depends on the questionable assumption that symptoms are expressed the same way in different cultures.
Methods: This comprehensive review used narrative synthesis to explore the psychometric, conceptual, and ethical aspects of the BDI studies done on adolescents in LMICs. We purposively selected 8 case exemplars from the literature based on a pre-specified framework to illustrate key cross-cultural and methodological issues. A custom rubric was used to weight studies by methodological quality. Invariance evidence was defined as explicit reporting of multi-group confirmatory factor analysis (MGCFA) for configural, metric, or scalar invariance.
Results: Although the BDI showed high internal consistency across the 8 exemplar studies (Cronbach's : 0.79–0.96), it revealed significant psychometric variation across cultures. Somatic items showed substantial overlap with medical comorbidities (e.g., HIV, malnutrition) in vulnerable populations, confounding diagnostic specificity. Factor structures were inconsistent, and rigorous measurement invariance testing was largely absent.
Discussion: The widespread absence of scalar invariance suggests that current cross-cultural score comparisons may reflect measurement artifacts rather than true differences in depression severity. Beyond psychometrics, BDI presents ethical challenges regarding developmentally inappropriate items, particularly concerning sexuality and suicidality, that limit its validity in conservative or resource-limited contexts.
Conclusion: Investigations into this issue should focus more on the use of invariance-testing frameworks, the implementation of a mixed-methods strategy, and the recognition of the need for sensitive translation and culturally appropriate assessment techniques.

1. Introduction

Depression among adolescents is a major worldwide public health problem. Unfortunately, the majority of cases occur in low- and middle-income countries (LMICs) that have very limited mental health resources [1, 2]. The Beck Depression Inventory (BDI-I and BDI-II), grounded in Beck's cognitive theory of depression, has become commonly used for depression symptom measurement worldwide [3]. Nonetheless, the extensive use of BDI in LMICs often follows what has been termed "methodological imperialism"—the practice of translating and validating Western-developed instruments without adequate consideration of their cultural and epistemological foundations [4, 5].

The main premise of cross-cultural psychological assessment, i.e. one tool can measure the same constructs in different populations, requires empirical proof of measurement invariance [6]. This challenge intensifies when assessing adolescents, whose emotional awareness and capacity for cognitive abstraction follow developmental trajectories that may interact with cultural context [7]. Furthermore, cultural syndromes and idioms of distress, particularly the somatization of psychological distress, which is strongly present in numerous non-Western cultural groups, challenge the BDI's cognitive-affective-somatic framework and point to the importance of cultural-contextual factors in adolescent mental health assessment [8].

Recent epidemiological data highlight the need for such an investigation. A systematic review of mental health problems among adolescents in sub-Saharan Africa revealed that 26.9% (IQR 20.1–31.1) of general population adolescents are showing depressive symptoms, and the percentage goes even higher in at-risk groups [9]. Correspondingly, approximately 24.4% (95% CI 19.2-30.5)

* Corresponding author: Anas Al-Qatefy, Faculty of Medicine, Mansoura National University, Mansoura, Egypt. Email: aanwar@std.mansnu.edu.eg

Published by the American Society for Inclusion, Diversity, and Equity in Healthcare (ASIDE). ISSN (Print) 3067-8730, ISSN (Online) 3067-8749. © 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Hosting by ASIDE Journals.

Citation: Al-Qatefy A, Tolba A, Eldawy Z, Ghallab M. Cross-Cultural Application of the Beck Depression Inventory (BDI-I/II) for Assessing Adolescent Depression in Low- and Middle-Income Countries: A Critical Review of Evidence from the Last Quarter-Century. ASIDE Health Sci. 2026;2(2):24-32, doi:10.71079/ASIDE.HS.012226419

of university students in LMICs were identified with depressive symptoms [10]. Almost 90% of adolescents globally reside in LMICs, meaning that even moderate prevalence translates into a massive public health burden [11].

Multiple attempts to validate the BDI have been made. Still, no comprehensive synthesis has yet psychometrically and conceptually examined the BDI (both versions, I and II) among adolescents in LMICs over the last 25 years. This critical narrative review builds on prior validation studies of the BDI with adolescent populations in LMICs over the last 25 years. We go beyond simply listing psychometric properties to re-evaluate the evidence for the BDI's cross-cultural construct validity, examine methodological rigor, interpret challenges in understanding conceptual equivalence, and reflect on the ethical and scientific aspects of global adolescent mental health assessment solutions that stem from this literature.

2. Methods

2.1. Review Design

The study employs a critical narrative review method, selected because it is appropriate for synthesizing heterogeneous literature and for offering a conceptual critique of intricate phenomena, according to Greenhalgh et al. [12]. Our synthesis includes studies using both the original BDI (BDI-I) and the revised BDI-II. While both versions assess depressive symptomatology, they differ substantively: BDI-I reflects symptoms over the past week and includes items on weight loss, somatic preoccupation, and work difficulty, whereas BDI-II assesses symptoms over the past two weeks and replaces several items (e.g., adding "changes in sleeping pattern" and "concentration difficulty") while removing others. These structural differences may independently influence factor structure, reliability, and optimal cut-offs, regardless of cultural context.

2.2. Search Strategy and Selection Criteria

A structured search was conducted in PubMed, PsycINFO, and Scopus covering January 2000 to January 2025. The search strategy, as elaborated in Supplementary File 1, employed: ("Beck Depression Inventory" OR "BDI") AND (adolescent* OR youth OR teenage*) AND (valid* OR reliab* OR psychometric*) AND (LMIC OR "low and middle income" OR "developing countr*" OR "global health"). Reference lists of assessed papers were manually screened to identify additional relevant studies. Inclusion criteria covered empirical studies that demonstrated the validation of the BDI-I or BDI-II in adolescent populations (ages 10–19) living in low- and middle-income countries (LMICs) and that reported at least one psychometric property (e.g., reliability, validity, factor structure). The status of lower- and middle-income countries was defined according to the World Bank Group's 2023 country income classification. Exclusion criteria were studies that were completely adult-only, non-English, or non-translated papers without an English abstract, and theoretical discussions.

2.3. Synthesis

Critical narrative reviews are well-suited to interrogating conceptual assumptions and synthesizing heterogeneous evidence through purposive selection of illustrative cases. A narrative synthesis was performed through thematic analysis organized around cultural, methodological, and developmental challenges. From the 54 full-text studies assessed for relevance, we purposively selected 8 case exemplars based on a pre-specified framework emphasizing: (1) geographic diversity across LMIC regions, (2) variation in cultural adaptation procedures, and (3) depth of psychometric reporting (particularly factor structure or measurement invariance). These exemplars were chosen to maximize insight into cross-cultural validity challenges rather than statistical generalizability. The remaining 46 studies informed contextual understanding but were excluded from detailed synthesis due to insufficient psychometric reporting, methodological limitations, or sample ineligibility, as elaborated in Supplementary File 2. Additionally, we applied a custom quality rubric assessing four criteria: (1) sample representativeness, (2) use of cognitive interviewing or local input in adaptation (beyond back-translation), (3) reliability reporting (e.g., Cronbach's), and (4) critical discussion of validity (e.g., item equivalence or cut-off appropriateness). Each criterion scored 1 (met) or 0 (not met); studies scoring 2 were selected as exemplars. Two reviewers appraised independently; disagreements were resolved by consensus. Invariance evidence was defined as explicit reporting of multi-group confirmatory factor analysis (MGCFA) for configural, metric, or scalar invariance; claims without statistical testing were not counted. Following Cohen's general effect size guidelines (often applied in psychological assessment), convergent validity coefficients 0.50 are interpreted as strong, 0.30–0.49 as moderate, and < 0.30 as weak [13].

3. Review

3.1. Overview of Included Studies

This narrative review summarized the results of eight representative validation studies of the Beck Depression Inventory (BDI-I and BDI-II) conducted in adolescent populations in low- and middle-income countries (LMICs), as shown in **(Table 1)**. Our exemplar studies span three distinct population contexts: (1) general adolescent populations (school-based samples in Nigeria, India, Brazil, Chile, Iran, and Turkey), (2) clinical or medically vulnerable groups (HIV-affected youth in Malawi [22]), and (3) institutionally at-risk cohorts (orphaned adolescents in Nepal [18]). These settings differ markedly in baseline distress levels, symptom expression, and comorbidity profiles. The studies included 181-1,412 adolescent participants, and in most cases, the adolescents ranged in age from 13 to 18 years. Various forms of BDI were also applied, including the original BDI, BDI-II in several translated versions, and culturally adapted versions. Translation-back-translation was a common protocol across all studies; however, the extent and quality of the cultural adaptation process varied significantly from one study to another.

3.1.1. Misleading Reliability Paradox

Empirical evidence from the included exemplar studies indicates that the Cronbach's alpha coefficients were high across all eight studies, ranging from 0.79 to 0.96. In India, the back-translated Tamil version of the BDI showed excellent internal consistency (= 0.96) and good 4-week test-retest reliability (r = 0.82). It also had high convergent validity with the Children's Depression Rating Scale-Revised (r = 0.72) [15]. In orphanage environments, the Nepali BDI-II translation had = 0.79 [18], while the Persian BDI-II in Iranian high school students showed = 0.83 and a test-retest reliability of r 0.55 [19]. Several studies on Brazilian adolescents reported good internal consistency across samples, and Ferreira-Maia and colleagues, in particular, reported elevated alpha values in their large São Paulo sample of 1,184 adolescents aged 13-18 years (= 0.88) [21].

These findings raise critical methodological implications: the consistent pattern of very high internal consistency across the board is quite impressive but may also be somewhat deceptive. The high alpha coefficients indicate that the items correlate strongly within each culture, i.e., respondents in each culture respond to the items in a statistically coherent way. Nevertheless, this single psychometric

Table 1: Cross-cultural validation studies of the Beck Depression Inventory (BDI) in adolescent populations

Author (Year)	Country / Region	BDI Version and Cultural Adaptation	Sample (n, age)	Factor Structure / Validity Evidence
(Adewuya et al., 2007) [14]	Nigeria (South-West)	BDI-I (Translation and back-translation)	n = 1,095; school adolescents aged 13–18 years	Criterion validation against K-SADS/E; ROC analyses confirmed strong diagnostic accuracy.
(Basker et al., 2007) [15]	India (Vellore, Tamil Nadu)	BDI-I (Tamil back-translation)	n = 181; secondary school adolescents	EFA and CFA replicated a 2-factor solution (30.5% variance explained); distinct screening and diagnostic cut-offs were proposed (5 vs 22).
(Araya et al., 2013) [16]	Chile (Santiago)	BDI-II (Spanish adaptation)	n = 592; mean age 15.5 years	EFA/CFA supported a unidimensional structure; criterion validity was established against MINI-KID; sex-specific cut-offs were identified (higher for girls).
(Kim et al., 2014) [17]	Malawi	BDI-II (HIV-affected cohort)	n = 562; adolescents aged 12–18 years in pediatric HIV program	Criterion validity confirmed; authors noted potential inflation of somatic symptoms due to HIV comorbidity. AUC = 0.82 (95% CI: 0.78–0.89) against CDRS-R as criterion.
(Bhatt et al., 2020) [18]	Nepal (Childcare homes)	BDI-II (Nepali version with pilot testing)	n = 602; adolescents aged 13–17 years	Internal consistency and convergent validity are reported; the authors caution against interpreting the results in the context of the study.
(Toosi et al., 2017) [19]	Iran (Shiraz)	BDI-II (Persian translation with expert review)	n = 1,150 (main sample); reliability subsample n = 128; high-school students	EFA (varimax) yielded a 4-factor solution, explaining 43.7% of the variance; concurrent validity with DASS-21 and BAI was confirmed.
(Jha et al., 2017) [20]	India (Bihar and multi-site school surveys)	BDI-II (Hindi translations)	In 1,412, school adolescents aged 14–18 years	Reported high prevalence estimates; methodological warnings against adult cut-off application to adolescents.
(Ferreira-Maia et al., 2018) [21]	Brazil (São Paulo)	BDI-II (Portuguese adaptation)	In 1,184, adolescents aged 13–18 years	EFA/CFA supported a dominant general factor; gender invariance analyses indicated minor differential item functioning.

BDI, Beck Depression Inventory; BDI-II, Beck Depression Inventory–Second Edition; , Cronbach's alpha; r, correlation coefficient; ROC, Receiver Operating Characteristic; Sn, Sensitivity; Sp, Specificity; K-SADS/E, Schedule for Affective Disorders and Schizophrenia for School-Age Children—Epidemiologic Version; CDRS-R, Children's Depression Rating Scale–Revised; MINI-KID, Mini International Neuropsychiatric Interview for Children and Adolescents; DASS-21, Depression Anxiety Stress Scales–21; BAI, Beck Anxiety Inventory; EFA, Exploratory Factor Analysis; CFA, Confirmatory Factor Analysis.

characteristic cannot establish the construct's equivalence across cultures, which is the most important issue [23]. Sijtsma (2009) argued that high internal consistency is a prerequisite but not sufficient for construct validity, especially in cross-cultural contexts, where high item correlations may be due to cultural factors rather than to shared cultural response biases, acquiescence, or socially desirable responding patterns [24, 25]. In other words, internal consistency supports within-group reliability but does not guarantee between-group comparability—a distinction often overlooked in cross-cultural psychometric research.

The Nigerian validation study conducted by Adewuya et al. offers valuable insights. The study demonstrated excellent screening performance, with a sensitivity of 0.91 and specificity of 0.97 at a cut-off of 18 on the Kiddie Schedule for Affective Disorders and Schizophrenia/Epidemiologic version (K-SADS/E) diagnostic interview. However, the authors strongly advised against the direct use of Western cut-offs, stressing the importance of local threshold validation and noting significant gender differences, with a higher prevalence in female adolescents. The finding that female adolescents are most affected underscores the critical methodological point: acceptable psychometric properties within a single cultural context do not automatically justify cross-cultural score comparisons or the application of universal cut-off values [14].

These reliability measures were almost entirely based on school-based convenience sampling of children recruited. Thus, the alpha coefficients may have been inflated due to restriction of range and spectrum bias [26]. Large samples of schools (n 1,412 in one Bihar sample) using Hindi translations of the BDI-II, as reported in the Indian school surveys by Jha and colleagues [27], showed high internal consistency but very high point prevalence estimates. Critics suggested that these prevalence figures were inflated because adult-derived Western cut-offs for adolescents were uncritically applied to populations without validation, thereby illustrating how psychometric properties that appear to be reliable can be used to conceal cross-cultural issues in instrument application.

Additionally, cross-cultural response styles—including extreme responding, midpoint avoidance, and socially desirable answering—may further inflate internal consistency and distort item means independently of depressive symptomatology. Yet, these patterns were rarely assessed in the included validations. The Chimera of

Conceptual Equivalence A major finding from the studies' review is that the research identified significant differences in the factor structures across different cultural settings, thereby endangering the conceptual equivalence assumption. To investigate this, Basker et al. conducted an Indian Tamil validation and used both exploratory and confirmatory factor analyses; the final two-factor solution they identified explained about 30.5% of the variance [15]. The study particularly distinguished between screening cut-offs (5) and diagnostic cut-offs (22), suggesting that the BDI could be effective in a clinical context where situations and severity levels change.

While in the Indian Tamil context the personality of BDI may differ, the Chilean study by Araya et al. found that exploratory and confirmatory analyses both supported a largely unidimensional solution among 592 adolescents (mean age 15.5 years) in Santiago [16]. This result could have been taken as evidence for the cross-cultural stability of the model if not for the study's disclosure of a critical complication: cut-offs for males and females differed substantially, with girls requiring higher thresholds than boys. To further support their point, the authors noted that using a single cut-off value derived from a Western sample can lead to errors when applied to other genders, and that a need for sex- and context-specific validation should be recognized instead. This sex-based differential functioning implies that the association between the observed and the latent variables of depression constructs may change for demographic subgroups even within a single cultural context.

The Persian BDI-II validation among Iranian high school students by Toosi et al. revealed yet another factor structure. The exploratory factor analysis with varimax rotation gave a four-factor solution accounting for approximately 43.7% of the variance. The authors concluded that the BDI-II could still be a viable instrument for measuring depression in Iranian teen populations; however, due to disparities at the item-level, changes were required in the cultural context, and conducting cognitive interviews was necessary, especially with younger adolescents, as they might understand the items in a different way than older ones [19].

The Brazilian study by Ferreira-Maia and colleagues indeed represents the most methodologically sophisticated examination of factor structure and measurement invariance among the reviewed studies. Their abstract and subsequent Brazilian adolescent research reported that exploratory and confirmatory factor analyses revealed a dominant general depression factor alongside reliable somatic and cognitive group factors. Critically, they tested measurement invariance across gender and documented gender differences in item functioning. This finding supports the presence of a strong general depression construct. Still, it demonstrates that the relationship between this latent construct and specific item responses varies by gender—a psychometric property that undermines simple score comparisons between male and female adolescents.

Interpretively, this heterogeneity challenges the assumption that the BDI measures a single, equivalent depression construct across cultures. These divergent factor structures—two-factor, unidimensional, four-factor, and hierarchical models—across cultural contexts are not merely technical statistical details. They reflect fundamental differences in how depressive symptoms cluster and co-occur in different populations. As Chen argued, "when factor structures differ across groups, the construct being measured may itself differ, rendering cross-group comparisons conceptually problematic" [28]. The BDI may be assessing somewhat different constellations of distress in Nigeria, India, Chile, Iran, and Brazil, even when items are competently translated.

Medical Comorbidity and the Somatic Symptom Confound Kim et al.'s paper on the Malawian HIV-affected adolescent population brings to light how medical comorbidity can weaken the discriminant validity of the BDI's somatic items. While the study reported acceptable internal consistency and concurrent validity against structured diagnostic scales (CDRS-R), the authors issued a critical caution: somatic items assessing fatigue, sleep disturbance, and appetite loss showed substantial overlap with HIV-related symptoms [22]. In contexts where adolescents suffer from chronic diseases, these items become powerless in terms of diagnostic specificity for depression.

The Nepali study done by Bhatt et al. concentrated on the performance of the BDI-II among 371 orphaned and vulnerable adolescents (13-17 years old) in childcare institutions. The authors emphasized the critical importance of accounting for trauma, loss, and institutional context when interpreting somatic and affective items. They recommended mixed-methods approaches for adaptation, recognizing that symptoms arising from grief, institutional deprivation, and trauma-related stress may be phenomenologically and etiologically distinct from symptoms of major depressive disorder, despite producing similar BDI score elevations [18].

This necessitates a reconceptualization of somatic items through the lens of "local biologies"– the acknowledgement that although human physiology is universal structurally, the manifestation and the meaning given to bodily experiences are very much dependent on the ecological, epidemiological, and social contexts [29]. In settings where malnutrition, parasitic infections, and infectious diseases are endemic, the phenomenological experience of "fatigue" or "appetite loss" fundamentally differs from experiences in resource-rich environments where these symptoms more reliably signal mood pathology [30]. The BDI is not capable of distinguishing fatigue caused by anemia, malaria, or intestinal parasites from the anhedonic fatigue that is typical of melancholic depression.

3.2. Criterion Validation and Cut-off Score Variability

Empirical evidence from exemplar studies demonstrates substantial variability in optimal cut-offs; several criterion-validation studies that underwent structured diagnostic interviews showed a wide range of optimal cut-offs. Adewuya et al., in their study in Nigeria, identified a cut-off of 18 as optimal for screening, yielding excellent sensitivity (0.91) and specificity (0.97) against K-SADS/E diagnoses. [14]. In contrast, the Indian Tamil validation by Basker et al. distinguished between a screening cut-off of 5 and a diagnostic cut-off of 22, suggesting that different thresholds may be appropriate depending on assessment purpose and clinical context [15].

This variability in cut-off results seriously challenges clinical practice and epidemiological research. Firstly, when researchers from various countries decide on different cut-offs to apply, or in case they uncritically apply a universal Western-derived threshold, cross-national prevalence estimates are not comparable. Secondly, the Chilean study's finding of sex-specific optimal cut-offs [16] Makes it even more difficult to deal with. If optimal thresholds vary not only across countries but also within each country by gender, this would suggest that population-level norms may be insufficient and that gender-specific reference data are needed.

3.3. The Neglected Adolescent Cognitive Context

Several studies noted developmental considerations that warrant attention. Toosi et al. even went so far as to suggest cognitive interviewing for younger adolescents, recognizing that early adolescents might interpret the abstract or metaphorical meanings of the items differently than older adolescents. The Basker et al. study, conducted

in primary care pediatric settings, emphasized the importance of pediatrician-administered support, suggesting that adolescents may need more structured support than is typically provided in self-administration.

These observations align with developmental neuroscience research documenting that adolescence involves ongoing maturation of prefrontal cortex regions responsible for abstract reasoning, emotional regulation, and metacognitive awareness [31]. Younger adolescents who are still mainly dependent on concrete operational thinking may interpret an item such as "I feel I am being punished" literally rather than metaphorically [32]. Similarly, items assessing self-evaluation and future orientation presuppose cognitive capacities for abstract self-reflection and temporal projection that continue to develop through adolescence [33]. Administration mode varied across exemplar studies and may contribute to response heterogeneity. Adewuya et al. [14] and most Brazilian/Chilean studies used standard self-administration. In contrast, Basker et al. [15] administered the BDI with pediatrician support in a clinical setting, while Kim et al. [22] and Bhatt et al. [18] used trained interviewers for orphaned or HIV-affected youth with potential literacy or trauma-related barriers. Clinician- or interviewer-assisted administration may reduce missing data and increase disclosure of sensitive items (e.g., suicidality, guilt), whereas self-report in school settings may heighten social desirability or literal interpretation—particularly in low-literacy contexts.

Considering that literacy and educational heterogeneity may further threaten construct validity in LMIC settings. The BDI assumes a minimum 5th–6th grade reading level in its original English form; translated versions rarely undergo readability validation. In contexts where adolescent literacy is variable—due to school dropout, limited educational access, or language mismatch—item comprehension may be compromised, biasing responses toward literal interpretation, acquiescence, or non-response. This risk is especially acute for abstract or metaphorical items (e.g., “I feel I am being punished”), which demand higher-order cognitive and linguistic processing [3].

Measurement Invariance and the Fallacy of Universal Cut-offs Perhaps the most critical methodological gap across the reviewed studies is the near-complete absence of measurement invariance testing, without which no valid cross-cultural score comparisons can be made. Among the eight exemplar studies, only the Chilean study by Araya et al. and the Brazilian study by Ferreira-Maia et al. explicitly referenced invariance testing—and even these established it only partially or noted gender-specific differences, which made interpretation difficult [16, 21]. The finding of Araya et al. that optimal cut-offs differ by sex is a direct indication that scalar invariance—the state where item intercepts are equal across groups—does not hold, because equal cut-offs would only be valid under scalar invariance.

This gap has profound psychometric consequences, as measurement invariance testing proceeds hierarchically through configural, metric, and scalar levels [34]. Configural invariance requires equivalent factor structure; metric invariance, equivalent factor loadings; and scalar invariance, equivalent item intercepts. Only after establishing scalar invariance can latent mean differences be interpreted as reflecting true differences in the underlying construct rather than measurement artifacts [28].

The lack of invariance testing represents a fundamental psychometric gap. A score of 18 on the BDI in a Nigerian, Chilean, Indian, or Iranian adolescent sample cannot be assumed to reflect equivalent depression severity if item intercepts differ due to cultural response styles, differential stigma, linguistic nuance, or divergent symptom conceptualization [35]. The use of Western-derived cut-offs in these circumstances conflates true disorder prevalence with measurement bias—a problem that is raised in Indian school prevalence studies [27].

Empirically, the failure to establish scalar invariance is mirrored in the divergent factor structures reported across validation studies [15, 19, 21], sex-specific optimal thresholds [16], and context-dependent somatic item interpretations [18, 22]. The findings, in fact, challenge the whole idea that BDI cut-off scores have the same diagnostic value across different cultural-linguistic settings. When scalar invariance is not achieved, the symptoms of two adolescents having the same level of depression -for example, one in Chile and the other in Nigeria- may reflect differently on their respective test scores because of culture-influenced ways of responding, stigma, or the translation meaning [36].

Epidemiological reports based on Western cut-offs in LMIC populations will likely yield an incorrect estimate of prevalence [24], while clinical screening or monitoring based on unvalidated thresholds risks systematic misclassification. Future validation research must therefore prioritize multi-group confirmatory factor analysis to sequentially test configural, metric, and scalar invariance across gender, age, and cultural subgroups [34, 37]. When invariance fails, researchers are required to identify items that function differentially, modify or discard them, or create context-specific norms indicative of culturally grounded symptom expression rather than the application of external psychometric universals. Cultural Psychology and the Limits of Beck’s Cognitive Model The poor and inconsistent performance of cognitive-affective items across multiple validation studies [14, 16, 19] reveals fundamental tension between the BDI’s theoretical foundation and non-Western constructions of selfhood. This reflects deeper ontological differences in selfhood; as Beck’s cognitive model presupposes an internalized, individualized self, in which negative automatic thoughts about personal worth, competence, and the future constitute core depressive pathology [38]. Yet, in collectivist cultures existing in numerous LMICs, the concept of self is more interdependent and is mainly determined by one’s social roles, family relationships, and community status rather than personal characteristics [39].

This ontological difference has profound implications for symptom expression. Chentsova-Dutton and Tsai demonstrate that cultural values fundamentally shape not merely how depression is described but how it is phenomenologically experienced [40]. In individualist cultures, depressed individuals emphasize personal failure, worthlessness, and internal negative self-evaluation—precisely the cognitive content the BDI was designed to assess. On the other hand, in collectivist cultures, individuals may feel and communicate their distress primarily through physical symptoms, disruptions in interpersonal relations, or spiritual aspects rather than self-criticism [41, 42].

The sex differences in emotional expressivity items reported by Adewuya et al., and the sex-specific cut-offs identified by Araya et al., likely reflect culturally prescribed emotional display rules that interact with gender socialization [14, 16]. In many cultures, masculine ideals emphasize emotional stoicism, making male adolescents less likely to acknowledge symptoms related to crying, sadness display, or emotional vulnerability even when experiencing significant depressive distress [43].

These patterns necessitate a fundamental shift from "imposed etics"—treating Western psychological constructs as universal standards—toward combined "emic-etic" approaches that integrate universal dimensions of negative affect with locally salient idioms

of distress [44, 45]. This might involve retaining certain BDI items that assess transculturally relevant symptoms, while adding or changing those culturally problematic cognitive items with the local expressions of distress that make sense in the particular culture [46].

Ethical Imperatives and Guidance Beyond Psychometric Adequacy Beyond psychometric limitations, ethical concerns require explicit addressing. Item 21 ("Loss of interest in sex") has been repeatedly identified as developmentally inappropriate and culturally problematic for adolescents, particularly in contexts where discussing sexuality is taboo, where many adolescents lack sexual experience, or where sexual activity before marriage is religiously proscribed [47–49]. This item routinely produces missing data, systematic under-reporting, and discomfort among respondents [50–52].

Similarly, Item 9 evaluating suicidal ideation raises ethical challenges in contexts where suicidality is stigmatized from a moral point of view or where talking about such thoughts may result in negative consequences. This demands ethically grounded instrument design: ethical principles for research require that instruments be not only psychometrically robust but also culturally sensitive and that they take into account participants' developmental stage and the social/environmental context [53]. To address these concerns, we recommend: (1) omit or reword Item 21 in developmentally or culturally inappropriate contexts; (2) administer Item 9 only with a safety protocol (trained staff, debriefing, and referral pathways); (3) pilot-test all items with local adolescents to assess acceptability and comprehension prior to use; and (4) tell participants they may skip any item—no coercion, no full-score pressure.

3.4. Methodological Recommendations

Advancing cross-cultural adolescent mental health research requires methodological rigor balanced with contextual feasibility and ethical sensitivity. We propose a three-component framework:

First, adopt an invariance-first approach as standard practice. Before any cross-group comparison, a multi-group confirmatory factor analysis should test configural, metric, and scalar invariance across gender, age, and cultural groups [28, 34]. If scalar invariance fails, the researchers are advised to highlight the items that display differential functioning and then decide whether to modify, eliminate, or create group-specific scoring algorithms.

Second, integrate mixed-methods approaches. Quantitative validation must be complemented by qualitative inquiry, particularly cognitive interviews and focus groups with adolescents from target populations [46]. Cognitive interviewing is an approach that can reveal semantic misunderstandings, identify culturally inappropriate content, and even detect developmentally challenging phrasing, all of which can be done before large-scale validation.

Third, incorporate criterion validity assessment. Even when gold-standard diagnostic interviews (K-SADS, MINI) are resource-intensive, conducting them with purposefully selected subsamples can establish locally appropriate cut-off scores. The Nigerian study's approach offers a feasible example: carrying out diagnostic interviews with extreme scorers to figure out the best sensitivity-specificity trade-offs [14].

Fourth, embrace decolonizing co-creative frameworks. Rather than merely adapting Western instruments, future research should involve local researchers and adolescents in the generation of items and pretesting [54]. Co-creative methods can initially involve questions that are directed towards adolescents and local mental health stakeholders, asking them how they identify, narrate, and comprehend depression in their context.

Fifth, as mobile health (mHealth) screening expands in LMICs [55], future validations must assess whether BDI's psychometric properties—established predominantly via paper-pencil—are preserved in digital formats, where interface design, literacy demands, and privacy perceptions may alter response behavior.

3.5. Limitations

This narrative review has several important limitations. First, our selective synthesis of eight exemplar studies, while appropriate for narrative review methodology, as Green et al. discussed [56], necessarily excludes potentially relevant validation research and cannot provide precise quantitative effect-size estimates that meta-analysis would afford. Second, the restriction to English-language publications constitutes a major limitation, likely biasing regional representation and underrepresenting culturally grounded validation work from non-Anglophone LMICs. Third, publication bias likely affects the literature we reviewed, as exemplar-based synthesis limits generalizability. Fourth, our synthesis did not stratify findings by urban/rural setting, socioeconomic status, or time of data collection—factors that may substantially influence symptom expression and instrument validity in rapidly changing LMIC contexts, and did not evaluate alternative instruments (e.g., PHQ-A, CES-D, or locally developed tools) that may offer superior cross-cultural validity or feasibility for adolescents in LMICs.. Additionally, although two reviewers independently appraised study quality using our rubric and resolved disagreements by consensus, the initial exclusion of 46 full-text studies from in-depth synthesis was not independently verified. To enhance auditability, we now provide a supplementary table listing all excluded studies with their rubric scores and exclusion reasons. These version-specific differences (between BDI-I and BDI-II) may also independently influence psychometric properties, potentially confounding cultural and instrument-level variance. Future systematic reviews with meta-analytic synthesis offer more detailed quantitative summaries by utilizing tools such as COSMIN [53].

4. Conclusion

The Beck Depression Inventory, while representing a historically important milestone in depression assessment, embodies assumptions about symptom expression, selfhood, and psychological distress that demonstrably do not generalize uniformly across all cultural contexts. Based on the exemplar studies and broader literature, our critical review suggests that the BDI is a tool that consistently ranks (individuals within specific cultural-linguistic groups), as shown by high internal consistency; however, current evidence is insufficient to confirm whether the BDI measures equivalent constructs across diverse cultural contexts, highlighting the urgent need for formal measurement invariance testing.

The consistent pattern of high internal reliability creates what we term a "reliability paradox"—a dangerous illusion of cross-cultural utility that masks profound threats to construct validity. Factor structure heterogeneity, somatic symptom confounding in medically compromised populations, the almost complete lack of formal testing for scalar invariance, sex-specific cut-off requirements, and the context-dependent interpretation of cognitive-affective items together point out the fundamental limits to the cross-cultural applicability of the BDI.

These limitations reflect deeper epistemological tensions between universalist assumptions inherent in Western psychiatric nosology and the cultural psychology recognition that cultural meaning systems, social structures, and local biologies fundamentally shape psychological experiences.

Ethically and accurately assessing adolescent depression in LMICs necessitates the abandonment of the administrative convenience of mere exported instruments in favor of contextually grounded, developmentally sensitive, and psychometrically sound tools that emerge from collaborative, co-creative processes. The continued use of Western-derived instruments without rigorous cross-cultural validation—particularly without measurement invariance testing and locally derived cut-offs—raises serious concerns about the validity of epidemiological estimates, the appropriateness of clinical thresholds, and the comparability of research findings across cultural contexts.

Conflicts of Interest

The authors declare that they have no competing interests.

Funding Source

No funding was received for this work.

Acknowledgment

None.

Institutional Review Board (IRB)

The institutional review board (IRB) approval was not required for this study.

Large Language Model

A locally deployed large language model (Qwen2-VL-72B-Instruct API) for personal research use was used exclusively to enhance language quality, coherence, and logical consistency in limited parts of the Review section. No confidential, proprietary, or participant-related information was entered into the model. All AI-assisted content was carefully reviewed and approved by the authors, who assume full responsibility for the accuracy, originality, and integrity of the manuscript.

Authors Contribution

AA conceptualized the topic, conducted the literature search, ZE interpreted the data, and AT and MG supervised the manuscript. All authors read and approved the final version of the manuscript.

Data Availability

All data generated or analyzed during this study are included in this published article and its reference list.

References

1. Kieling C, Baker-Henningham H, Belfer M, Conti G, Ertem I, Omigbodun O, et al. Child and adolescent mental health worldwide: evidence for action. Lancet. 2011;378(9801):1515-25. [PMID: 22008427, https://doi.org/10.1016/S0140-6736(11)60827-1].
2. WHO. Mental health of adolescents [Internet]; 2025. Available from: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health.
3. Beck AT, Steer RA, Brown G. Beck Depression Inventory–II. PsycTESTS Dataset. 2011. [https://doi.org/10.1037/t00742-000].
4. Fernando S. Race and culture issues in mental health and some thoughts on ethnic identity. Counselling Psychology Quarterly. 2012;25(2):113-23. [https://doi.org/10.1080/09515070.2012.674299].
5. van de Vijver F, Tanzer NK. Bias and equivalence in cross-cultural assessment: an overview. European Review of Applied Psychology. 2004. [https://doi.org/10.1016/j.erap.2003.12.004].
6. Byrne BM, van de Vijver FJR. The maximum likelihood alignment approach to testing for approximate measurement invariance: A paradigmatic cross-cultural application. Psicothema. 2017;29(4):539-51. [PMID: 29048316, https://doi.org/10.7334/psicothema2017.178].
7. Crone EA, Dahl RE. Understanding adolescence as a period of social-affective engagement and goal flexibility. Nat Rev Neurosci. 2012;13(9):636-50. [PMID: 22903221, https://doi.org/10.1038/nrn3313].
8. Kirmayer LJ. Cultural variations in the clinical presentation of depression and anxiety: implications for diagnosis and treatment. J Clin Psychiatry. 2001;62 Suppl 13:22-8; discussion 29-30. [PMID: 11434415].
9. Jorns-Presentati A, Napp AK, Dessauvagie AS, Stein DJ, Jonker D, Breet E, et al. The prevalence of mental health problems in sub-Saharan adolescents: A systematic review. PLoS One. 2021;16(5):e0251689. [PMID: 33989357, PMCID: PMC8121357, https://doi.org/10.1371/journal.pone.0251689].
10. Akhtar P, Ma L, Waqas A, Naveed S, Li Y, Rahman A, et al. Prevalence of depression among university students in low and middle income countries (LMICs): a systematic review and meta-analysis. J Affect Disord. 2020;274:911-9. [PMID: 32664032, https://doi.org/10.1016/j.jad.2020.03.183].
11. Wubshet I, Engdawork K, Gebremarim S, Kanazayire C, Abbott P. The Social Determinants of Depression among Adolescents in Low- and Middle-Income Countries: A Scoping Review. https://doiorg/101101/2025040725325376. 2025. [https://doi.org/10.1101/2025.04.07.25325376].
12. Greenhalgh T, Thorne S, Malterud K. Time to challenge the spurious hierarchy of systematic over narrative reviews? Eur J Clin Invest. 2018;48(6):e12931. [PMID: 29578574, PMCID: PMC6001568, https://doi.org/10.1111/eci.12931].
13. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Routledge. 2013. [https://doi.org/10.4324/9780203771587].
14. Adewuya AO, Ola BA, Aloba OO. Prevalence of major depressive disorders and a validation of the Beck Depression Inventory among Nigerian adolescents. Eur Child Adolesc Psychiatry. 2007;16(5):287-92. [PMID: 17473949, https://doi.org/10.1007/s00787-006-0557-0].
15. Basker M, Moses PD, Russell S, Russell PS. The psychometric properties of Beck Depression Inventory for adolescent depression in a primary-care paediatric setting in India. Child Adolesc Psychiatry Ment Health. 2007;1(1):8. [PMID: 17688697, PMCID: PMC1976414, https://doi.org/10.1186/1753-2000-1-8].
16. Araya R, Montero-Marin J, Barroilhet S, Fritsch R, Gaete J, Montgomery A. Detecting depression among adolescents in Santiago, Chile: sex differences. BMC Psychiatry. 2013;13(1):122. [PMID: 23617306, PMCID: PMC3637822, https://doi.org/10.1186/1471-244X-13-122].
17. Kim MH, Mazenga AC, Devandra A, Ahmed S, Kazembe PN, Yu X, et al. Prevalence of depression and validation of the Beck Depression Inventory-II and the Children's Depression Inventory-Short amongst HIV-positive adolescents in Malawi. Journal of the International AIDS Society. 2014;17(1). [https://doi.org/10.7448/ias.17.1.18965].
18. Bhatt KB, Apidechkul T, Srichan P, Bhatt N. Depressive symptoms among orphans and vulnerable adolescents in childcare homes in Nepal: a cross-sectional study. BMC Psychiatry. 2020;20(1):466. [PMID: 32977787, PMCID: PMC7517808, https://doi.org/10.1186/s12888-020-02863-y].
19. Toosi F, Rahimi C, Sajjadi S. Psychometric Properties of Beck Depression Inventory-II for High School Children in Shiraz City, Iran. International Journal of School Health. 2017;In press(In press). [https://doi.org/10.5812/intjsh.41069].
20. Jha KK, Singh SK, Nirala SK, Kumar C, Kumar P, Aggrawal N. Prevalence of Depression among School-going Adolescents in an Urban Area of Bihar, India. Indian Journal of Psychological Medicine. 2017;39(3):287-92. [https://doi.org/10.4103/0253-7176.207326].

21. Ferreira-Maia AP, Boronat AC, Gorenstein C, Wang YP. 5.39 Reliability and Dimensional Validity of the Beck Depression Inventory-II (BDI-II) Among Brazilian Adolescents. Journal of the American Academy of Child & Adolescent Psychiatry. 2018;57(10):S239. [https://doi.org/10.1016/j.jaac.2018.09.332].
22. Kim MH, Mazenga AC, Devandra A, Ahmed S, Kazembe PN, Yu X, et al. Prevalence of depression and validation of the Beck Depression Inventory-II and the Children's Depression Inventory-Short amongst HIV-positive adolescents in Malawi. J Int AIDS Soc. 2014;17(1):18965. [PMID: 25085002, PMCID: PMC4119288, https://doi.org/10.7448/IAS.17.1.18965 18965].
23. Flake JK, Pek J, Hehman E. Construct Validation in Social and Personality Research. Social Psychological and Personality Science. 2017;8(4):370-8. [https://doi.org/10.1177/1948550617693063].
24. Bolton P, Neugebauer R, Ndogoni L. Prevalence of depression in rural Rwanda based on symptom and functional criteria. J Nerv Ment Dis. 2002;190(9):631-7. [PMID: 12357098, https://doi.org/10.1097/00005053-200209000-00009].
25. Sijtsma K. On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha. Psychometrika. 2009;74(1):107-20. [PMID: 20037639, PMCID: PMC2792363, https://doi.org/10.1007/s11336-008-9101-0].
26. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. Continuing Education in Anaesthesia Critical Care & Pain. 2008;8(6):221-3. [https://doi.org/10.1093/bjaceaccp/mkn041].
27. Jha KK, Singh SK, Nirala SK, Kumar C, Kumar P, Aggrawal N. Prevalence of Depression among School-going Adolescents in an Urban Area of Bihar, India. Indian J Psychol Med. 2017;39(3):287-92. [PMID: 28615762, PMCID: PMC5461838, https://doi.org/10.4103/0253-7176.207326].
28. Chen FF. What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. J Pers Soc Psychol. 2008;95(5):1005-18. [PMID: 18954190, https://doi.org/10.1037/a0013193].
29. Lock M. Cultivating the Body: Anthropology and Epistemologies of Bodily Practice and Knowledge. Annual Review of Anthropology. 1993;22(1):133-55. [https://doi.org/10.1146/annurev.an.22.100193.001025].
30. Kohrt BA, Hruschka DJ. Nepali concepts of psychological trauma: the role of idioms of distress, ethnopsychology and ethnophysiology in alleviating suffering and preventing stigma. Cult Med Psychiatry. 2010;34(2):322-52. [PMID: 20309724, PMCID: PMC3819627, https://doi.org/10.1007/s11013-010-9170-2].
31. Blakemore SJ, Mills KL. Is adolescence a sensitive period for sociocultural processing? Annu Rev Psychol. 2014;65(1):187-207. [PMID: 24016274, https://doi.org/10.1146/annurev-psych-010213-115202].
32. Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, Angell KE. Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. J Abnorm Psychol. 1998;107(1):128-40. [PMID: 9505045, https://doi.org/10.1037//0021-843x.107.1.128].
33. Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. Lancet. 2012;379(9820):1056-67. [PMID: 22305766, PMCID: PMC3488279, https://doi.org/10.1016/S0140-6736(11)60871-4].
34. Putnick DL, Bornstein MH. Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. Dev Rev. 2016;41:71-90. [PMID: 27942093, PMCID: PMC5145197, https://doi.org/10.1016/j.dr.2016.06.004].
35. Gregorich SE. Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. Med Care. 2006;44(11 Suppl 3):S78-94. [PMID: 17060839, PMCID: PMC1808350, https://doi.org/10.1097/01.mlr.0000245454.12228.8f].
36. Weiss B, Dang M, Trung L, Nguyen MC, Thuy NT, Pollack A. A Nationally-Representative Epidemiological and Risk Factor Assessment of Child Mental Health in Vietnam. Int Perspect Psychol. 2014;3(3):139-53. [PMID: 25328817, PMCID: PMC4199236, https://doi.org/10.1037/ipp0000016].
37. Davidov E, Meuleman B, Cieciuch J, Schmidt P, Billiet J. Measurement Equivalence in Cross-National Research. Annual Review of Sociology. 2014;40(1):55-75. [https://doi.org/10.1146/annurev-soc-071913-043137].
38. Beck AT. The evolution of the cognitive model of depression and its neurobiological correlates. Am J Psychiatry. 2008;165(8):969-77. [PMID: 18628348, https://doi.org/10.1176/appi.ajp.2008.08050721].
39. Markus HR, Kitayama S. Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review. 1991;98(2):224-53. [https://doi.org/10.1037/0033-295x.98.2.224].
40. Chentsova-Dutton YE, Tsai JL. Self-focused attention and emotional reactivity: the role of culture. J Pers Soc Psychol. 2010;98(3):507-19. [PMID: 20175627, https://doi.org/10.1037/a0018534].
41. Kirmayer LJ, Young A. Culture and somatization: clinical, epidemiological, and ethnographic perspectives. Psychosom Med. 1998;60(4):420-30. [PMID: 9710287, https://doi.org/10.1097/00006842-199807000-00006].
42. Ryder AG, Yang J, Zhu X, Yao S, Yi J, Heine SJ, et al. The cultural shaping of depression: somatic symptoms in China, psychological symptoms in North America? J Abnorm Psychol. 2008;117(2):300-13. [PMID: 18489206, https://doi.org/10.1037/0021-843X.117.2.300].
43. Chaplin TM, Aldao A. Gender differences in emotion expression in children: a meta-analytic review. Psychol Bull. 2013;139(4):735-65. [PMID: 23231534, PMCID: PMC3597769, https://doi.org/10.1037/a0030737].
44. Canino G, Alegria M. Psychiatric diagnosis - is it universal or relative to culture? J Child Psychol Psychiatry. 2008;49(3):237-50. [PMID: 18333929, PMCID: PMC3104469, https://doi.org/10.1111/j.1469-7610.2007.01854.x].
45. Kleinman AM. Depression, somatization and the "new cross-cultural psychiatry". Soc Sci Med (1967). 1977;11(1):3-10. [PMID: 887955, https://doi.org/10.1016/0037-7856(77)90138-x].
46. Kaiser BN, Ticao C, Anoje C, Minto J, Boglosa J, Kohrt BA. Adapting culturally appropriate mental health screening tools for use among conflict-affected and other vulnerable adolescents in Nigeria. Glob Ment Health (Camb). 2019;6:e10. [PMID: 31258924, PMCID: PMC6582460, https://doi.org/10.1017/gmh.2019.8].
47. Askari F, Mirzaiinajmabadi K, Saeedy Rezvani M, Asgharinekah SM. Sexual health education issues (challenges) for adolescent boys in Iran: A qualitative study. J Educ Health Promot. 2020;9(1):33. [PMID: 32318601, PMCID: PMC7161661, https://doi.org/10.4103/jehp.jehp_462_19].
48. Joodaki K, Nedjat S, Vahid Dastjerdi M, Larijani B. Ethical considerations and challenges of sex education for adolescents in Iran: a qualitative study. J Med Ethics Hist Med. 2020;13:2. [PMID: 33088429, PMCID: PMC7569533, https://doi.org/10.18502/jmehm.v13i2.2664].
49. Osman A, Barrios FX, Gutierrez PM, Williams JE, Bailey J. Psychometric properties of the Beck Depression Inventory-II in nonclinical adolescent samples. J Clin Psychol. 2008;64(1):83-102. [PMID: 18161034, https://doi.org/10.1002/jclp.20433].
50. Lipton MF, Augenstein TM, Weeks JW, De Los Reyes A. A Multi-informant Approach to Assessing Fear of Positive Evaluation in Socially Anxious Adolescents. Journal of Child and Family Studies. 2013;23(7):1247-57. [https://doi.org/10.1007/s10826-013-9785-3].
51. De Los Reyes A, Makol BA, Racz SJ, Youngstrom EA, Lerner MD, Keeley LM. The Work and Social Adjustment Scale for Youth: A Measure for Assessing Youth Psychosocial Impairment Regardless of Mental Health Status. J Child Fam Stud. 2019;28(1):1-16. [PMID: 33311964, PMCID: PMC7731438, https://doi.org/10.1007/s10826-018-1238-6].
52. Rausch E, Racz SJ, Augenstein TM, Keeley L, Lipton MF, Szollos S, et al. A Multi-Informant Approach to Measuring Depressive Symptoms in Clinical Assessments of Adolescent Social Anxiety Using the Beck Depression Inventory-II: Convergent, Incremental, and Criterion-Related Validity. Child & Youth Care Forum. 2017;46(5):661-83. [https://doi.org/10.1007/s10566-017-9403-4].
53. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147-57. [PMID: 29435801, PMCID: PMC5891568, https://doi.org/10.1007/s11136-018-1798-3].

54. Lewis-Fernández R, Aggarwal NK, Hinton L, Hinton DE, Kirmayer LJ. DSM-5® Handbook on the Cultural Formulation Interview. DSM-5® Handbook on the Cultural Formulation Interview. 2015. [https://doi.org/10.1176/appi.books.9781615373567].
55. Bhatia RK, Bhardwaj P, Chaudhry A, Ghosh A, Parmar A, Ahluwalia Y, et al. A review of systematic reviews for evidence on use of mobile applications for mental disorders, including substance use disorders, in low and middle-income countries. Digit Health. 2025;11:20552076251387056. [PMID: 41143094, PMCID: PMC12547136, https://doi.org/10.1177/20552076251387056].
56. Green BN, Johnson CD, Adams A. Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. J Chiropr Med. 2006;5(3):101-17. [PMID: 19674681, PMCID: PMC2647067, https://doi.org/10.1016/S0899-3467(07)60142-6].