Development, Validation and Application of Becks Depression Inventory – 11 (BDI) In the Kenyan Context

Florence Adhiambo Ochanda., Dr Jasper Isoe,

Tangaza University Nairobi

DOI: https://dx.doi.org/10.47772/IJRISS.2025.910000747

Received: 02 November 2025; Accepted: 10 November 2025; Published: 22 November 2025

ABSTRACT

The paper reviewed the development, Validation, and Application of the BDI-11 in the Kenyan context. The review established the depth to which the BDI-11 tool had been validated and used within the local clinical and research setting. Peer reviewed studies and theses were revised, and those published between 2016 and 2022 were studied. From the studies the findings revealed that majority of the studies used the original western version which had not been adapted thus exhibiting low linguistic and cultural adaptation within the Kenyan context. The BDI-11 was validated with reliability demonstrated and the three-factor structure meeting the global findings, it was observed that the cut off score variations and language into the local context of Kenya still remained a challenge. For local psychometric framework to improve diagnostic accuracy, there is need for large scale studies to be carried out and this remains a gap. The conclusion for this review is that the BDI-11 is a useful and valid tool to be used as a screening and severity assessment instrument. There is need for further validation and adaptation to ensure context relevance, alignment with cultural sensitivity and policy relevance in assessment and counselling practice for mental health services in Kenya. In addition, there is need to reinforce the work that has started and develop translated versions for use within the Kenyan context with the aim of establishing Kenyan context specific norms or cut-offs. The use of valid tools in the local context ensures that the mental health frequency and problems are truly captured and reflected hence informing early interventions, tracking on treatment outcomes over time and effective clinical management of clients.

Key Words: Becks Depression Inventory (BDI-11), Depression, Psychometric Validation, Cultural Adaptation, Kenyan Population and Mental Health Assessment

INTRODUCTION

Depression has become a public health concern in the recent years (Jorm & Muller, 2022) both locally and worldwide, affecting the emotional, social and occupational functioning. Effective diagnosis, treatment and evaluation is heavily dependant on accurate assessment of the depressive symptoms. The burden of depression therefore necessitates the value and importance of accurate tools that can capture severity across different populations. If not properly diagnosed depression contributes to people suffering. The instrument widely used for measuring depression is the Becks Depression Inventory, 2nd edition (BDI-11 is commonly cited in literature (Seppanen, 20220). According to Beck, Steer, and Brown (1996), the instrument was reformulated and developers of the Becks Depression Inventory specify that this is a self report instrument that has the capacity to measure the severity of depression symptoms. The BDI-11 has been revised and has 21 items that assesses cognitive, affective and somatic symptoms and aligns with the (American Psychiatric Association, 2022) DSM-IV-TR criteria for major depressive disorder, the items are answered on a four-point scale on the depression symptoms. BDI-11 has been shown to be reliable in community and large samples considering it is self administered and scoring process is an easy procedure. According to (Hasan et al., 2025), the instrument exhibits high reliable and internal consistency measures. Extensive research has pointed to the use of BDI-11 (Almeida et al., 2022) because of its psychometric properties robustness hence has been widely accepted across the world as a self report tool for screening depressive symptoms across varied populations. However, there is need for local validation due to the variations in cultural expression of depression symptoms. The review hence focuses on development, validation and application of the BDI-11 in the Kenyan context.

Background to Psychometrics in Counselling Psychology

In the late 19th and 20th centaury the use of psychometrics in counselling psychology emerged with the need to measure psychological attributes objectively, in a way that does not influence a person's feelings or opinions. The early founders included Galton (1883) who measured individual differences. Binet (1905) introduced the first intelligence test to his population. Spearman (1904) introduced the classical test theory and factor analysis, which was followed by Terman (1916) and Weschler (1939) who introduced the standardized intelligence scales. The tool underwent revisions between 1978 (Beck et al. 1979) and again in 1964 (Beck et al., 1996) hence BDI-1A and BDI-11 with the BDI-11 revised to measure severe depression especially for patients requiring medical inpatient care. Studies typically report reliability of the instrument.

Psychometrics and its role in assessment and intervention

Psychometrics belongs to a branch of psychology which deals with theory and methods of psychological measurement. Psychometrics is used to define and measure constructs and applied through the use of psychometric instruments with the aim of improving and understanding the well being of an individual (Costa, 2021). Psychometrics is also concerned with measurement and ability to expect behaviour, and even psychological traits with the aim of improving well being (Wolfgang et al., 2022).

Counsellors are able to use psychometrics tests to help develop case formulation which allows the counsellor to understand some of the psychological challenges or traits a client might be facing, for example understanding the severity of the symptom or even evaluate coping skills with the aim of treatment (Zufferey et al., 2022).

Clients in counselling are usually looking to get better and improve their quality of life through engaging counselling services, the use of sound psychometric tools can helps assess therapist capacity hence improving the value of counselling (Gori, 2022). In his work Lutz et al. (2022) emphasised the importance of counsellor psychometric monitoring with the aim of making efficacious decision making in addressing patient psychometric requirements and outcomes. He reinforces the value of continuous patient monitoring with the aim of improving patient outcomes. The field of counselling psychology is highly dependant on research evidence measuring constructs and validating new tools and looking at individual difference (Wisjen, 2022) with the aim improving the mental health of a client. Barkham et al. (2023) indicates the use of Routine Outcome Monitoring (ROM) that supports early detection and this therefore contributes to treatment efficacy and outcomes.

In conclusion, many studies have pointed to the importance of using reliable and valid methods while working with clients. This therefore contributes to continuously assessing and monitoring client therapeutic progress which eventually informs the counsellor decision making. There is also the aspect of using instruments that align with the context and culture thus ensuring that interventions are client cantered and meet the needs of the diverse population. In Kenya the uptake of counselling services is increasingly becoming accepted even though evaluation frameworks remain critical in advancing effectiveness in delivering interventions in the mental health sector. The BDI-11 is a useful instrument for assessment and diagnosis of depression.

Importance of Evidence-Based Assessment in Kenya Amidst Rising Mental Health

The global challenges in Kenya have exacerbated the mental health difficulties faced by individuals. Amidst a declining economy, increase in violence, family conflicts, divorce rates and not limited to physical illness, this has contributed to the mental fragility of the human person. According to the Kenya Demographic and Health Survey 2022 (KDHS) the prevalence at 3.84% of depression and anxiety has increased among adults between 15-49 years hence requiring professionals to screen, access and diagnose depression and other mental health disorders. The ethical code of KCPA (2025), indicates that while professionals are working with clients it is important to do no harm at whatever cost. The use of assessment tools that are not reliable and appropriate to the population and setting, can be considered harm to the patients, and especially if the tool has not been validated in the local context, specifically Kenya. According to Watson et al. (2022) the use of BDI-11 in a Kenyan postnatal sample population, has contributed to the cultural relevance and validity because the items translate well to the local context and population used upon. The instrument was able to discriminate between cases and those presenting as non-cases indicating good psychometric properties. The Kenya Mental Health Policy 2015-

2030, (Ministry of Health, 2015) has supported the need to embrace the use of decision making that is hinged on policy. Generating evidence from localised BDI-11 instrument can inform mental health strategies hence prioritize budgets that will prioritize mental health programs within the community placing emphasis on data driven planning.

Importance of BDI for counselling psychology in Kenya

For the longest time, psychologists in Kenya have had to rely on tools developed and validated from different contexts. For example, the BDI was not developed with the local Kenyan population in mind, this therefore means that measuring certain cultural expressions may risk wrong interpretation. Words expressions such as distress may become very difficult to interpret, the ability to interrogate one tool thoroughly well gives an opportunity to determine whether the items are applicable or meaningful and understood by the local population. Hence validating a tool means that the local realities are taken into consideration as their meanings align. According to (Abubakar et al., 2016) he found out that several items required adaptation to the cultural context to maintain conceptual evidence. Focusing on the scale in depth would also lead to analysing validity, psychometric reliability and taking into consideration the Kenyan sample. Thus, the tool will measure depression rather other parameters. The lack of validation might not give the true results impacting negatively on policy decisions, diagnosis and treatment outcomes.

Once a tool has been validated, it offers a baseline measure for other comparisons within the field. Population comparison changes over a period of time hence strengthening localised evidence-based interventions becomes critical. Ethical considerations focus on having trained professionals in the Kenyan context, by validating one tool professionally healthcare workers who are not specialised can be trained to screen and monitor depression even in remote areas with low budgets hence increasing use of verified instruments. In 2000 the tool was revised to accommodate cross cultural adaptations with the population being used among both clinical and non clinical majorly in Europe, Asia and Africa.

Development of the Scale

History of its Development

The original creator of Becks Depression Inventory (BDI, 1961) was a gentle man called (Beck et al., 1961) who was a psychiatrist researcher at the University of Pennsylvania in that year with the aim of measuring depression using standardized tool, not depending on clinical impressions. It was to be used within the medical and psychiatric setting and would also be used 1for accessing efficacy of interventions within psychotherapy and use of medication.

Beck et al. (1996) revised the tool to align with the DSM-IV for Major Depressive Disorder to improve the tool by removing items such as (weight loss and body image) and adding agitation and weight loss (Beck et al., 1996).

Beck Depression Inventory-11

The Beck Depression Inventory (BDI) Beck et al. (1961) Beck Depression Inventory -11(BDI-11) is an instrument used to measure severity of depression symptoms for children and adults from 13 years and above. The instrument assess depression and evaluates for the intensity of depression symptoms; Specifically, it screens for depression; Assess severity, monitor changes; support clinical diagnosis and provide data for research (Abubakar et al., 2016). The release of DSM-111-R and DSM-IV necessitated the revision and developing (BDI-11; (Beck et al., 1996), many researchers have reviewed the instrument and found it to have good psychometric properties, Almeida, et al. (2023), using BDI-11 Portuguese version found acceptable fit indices with three bi-factor models, Afonso et al. (2020) found the instrument to have Cronbach's alpha coefficient at α=0.89 (95% CI, 0.89-0.9). The psychometric properties of the BDI-11 have been studied across different groups and cultures not always in the same representative samples hence indicating cultural invariance, henceforth the need to examine the invariance before use in new settings (Seppanen et al., 2022) with the Cronbach's alpha ranging between α=0.87 and α=0.93 sighting adequate internal consistency. Another global study by do Nascimento, et

al. (2023) also indicated cultural invariance when using the tool, however the instrument indicated good reliability.

Equally confirmatory factor Analysis (CFA), (Cognitive-affective + somatic) of the Bangla-11(Hasan et al., 2025), indicated an acceptable fit among undergraduate students CFI at 0.955, RMSEA =0.042 and internal consistency α = 0.88 for cognitive-affective and Cronbach's alpha for somatic at α=0.73 respectively (Hasan et al., 2025). Even though, the result indicted support for the two-factor model it remained lower in the somatic factor indicating deviation from the norm of the model. At the regional level Ogakwu et al. (2022), reviewed the psychometric properties among a sample of 363 administrators in a Nigerian Secondary school and indicated that the psychometric properties were robust with good internal consistency; confirmatory factor analysis (CFA) reported at (0.973, RMSEA = 0.048) for the factor model.

Theoretical Framework Underpinning BDI

The theoretical framework underpinning this instrument was Cognitive Theory of Depression Theory (Beck et al., 1961; Beck, 1967). The theory is based on the understanding of psychological processes that underlie depression and is also the basis for the BDI content and structure. From Becks perspective depression not only comes as a result of conflicts and biological disturbances; he indicated that negative thinking patterns are the core of depression, distorted thinking patterns and dysfunctional beliefs (Beck, 1967) are the main of aspect of depressive experiences. Individuals who are struggling with depression often present with negative thinking patterns (Beck et al, 1961). The essentials of cognitive triad (Beck, 1976; Smith et al., 2022) which includes how thinking and feelings are interconnected to elicit a particular behaviour hence individuals have a negative view of themselves, the world and the future. Individuals are also predisposed to cognitive distortions which look at the role of thinking errors that reinforce issues such as depression (Beck, 1976). The components include thinking in absolute standings, which help shape the individuals view of the world leading to maintained negative emotions. The theory was developed with self reported indicators which are measurable domains cognitive (thinking), affective (emotional), somatic (physical) and the motivational (Beck, 1967).

Therefore, the BDI-11 measures symptoms associated with the effective, cognitive and somatic (Almedia et al., 2022). Because it measures the latent symptoms it has the ability to separate cognitive affective and somatic vegetative (Almeida et al., 2021). The cognitive mode of assessing depression as a model has been supported and widely appreciated (MacCowan, 2022). The cognitive theory is relevant (Krystina et al, 2024), appreciates the use of this model as the foundation of measuring cognitive components as key factors- in predicting depression. The cognitive triad (Kristina et al., 2024) indicates that schema errors and cognitive processing are key factors towards depression research. Almeida et al. (2022) found the that the BDI-11 model had three factors (cognitive, affective and somatic).

Target Population

Having been developed by (Beck, et al., 1961), the instrument was developed and was meant to be used for psychiatric outpatient and psychiatric inpatients who were especially being treated for depression having received a psychiatric diagnosis. This was a self reported measurement instrument with the aim of reporting the severity of depression with the focus being to measure intensity and not make a diagnosis on the clinical patients. The population for this assessment were English speaking Americans who were predominantly middle-class.

The tool has been used with different populations (Nzangi et al., 2022) used the tool among adolescents aged between 14 - 21 within selected public secondary schools to assess the prevalence of depression and found the prevalence at 58.9%. This indicates that the tool can be used among the school population. Mbithi et al. (2023) also researched among the adolescent population within the school and community samples citing them as having behavioural problems and higher anxiety levels with schools providing access to non-practical samples with the results indicating prevalence of anxiety and depression at 29.0% and 19.3%, this study being consistent with other studies in the context of the pandemic.

Angachi et al., (2022) based their research on the prevalence of anxiety and depression among cervical cancer patients in referral hospitals in Western, Kenya. This indicates that the tool could also be used in the clinical

setting with the prevalence of depression at 42.7% among the age of 40-49 years, primary level education at 42.2% and the married participants at 42,7% (Angachi et al., 2022). The BDI-11 yielded the depression level at 67% indicating efficacy of the tool.

The BDI-11 has also been used among university and emerging adults within the university settings (Nteere et al., 2016), as these are common settings for specific stressors such as substance use, academic pressure and not limited to transition periods. The study results indicate Cronbach's alpha ranging from α = .82-.92 among Kenyan and South African samples (Osborn et al., 202). This implies that BdI-11 items measure common constructs and are generally cohesive. The two-factor mode yielded (cognitive-affective and somatic components, convergent validity r = .63-.72 supporting the argument that BDI-11 measures depression with other validated tools consistently. On the cut-off scores (Nteere et al., 2016) noted that the western-cut-off values had overestimated values for depression in the Kenyan samples, due to somatic loading and cultural interpretation thus a slightly higher threshold recommended for moderate to severe depression classification in Kenya. Mean differences were equally noted between clinical participants presenting higher scores while non clinical samples presenting with lower scores (Nteere et al., 2016).

Abubakar et al. (2016) in his research inventory among the low literacy population in the context of HIV used the BDI – 11, this pointed out the importance of a scale having the ability to work across both rural and urban area settings, socio-economic status and literacy level across a diverse population affecting comprehension and symptom expression of items. The key findings in his research pointed to the need for translation and interviewer administration as necessary where literacy levels are considered low and translation to accommodate local idioms and translation of the instrument (Abubakar et al., 2016).

Constructs

The inventory is a self report measure inventory covering different domains such as mood (sadness and pessimism), Cognition (negative thoughts, self-criticism), somatic/affective (sleep disturbance, appetite and fatigue) (Beck et al., 1961) (self-dislike, guilt and hopelessness, Motivation and loss of interest (Beck, et al., 1961). The revised version (Beck et al., 1996) improved to measuring Cognitive-Affective and Somatic-Performance dimensions, this aligned with diagnostic criterion in the (DSM IVTR 2022).

Other researchers (Almeida et al., 2022) measured Cognitive, Affective and somatic components of depression while using the two-factor model. (Htut et al., 2022) evaluated the cognitive, Affective and somatic dimensions using the three-factor model in the Mynmar clinical population of those who abused substances. Sanchez et al., (2022) in assessing factor structure and normative data indicated that the tool had good psychometric properties and was appropriate for their context and population with the total score not generalized but sufficient in some contexts.

Validation of the Scale

Global Validation

The use of BDI-11 has been approved and validated in many contexts and different populations globally. This is a self-report scale which measures the presence and severity of depressive symptoms across many countries. Seppanen et al. (2022) compared item level scores across six population samples in different countries and found the internal consistency to be Cronbach's α to range between 0.87 to 0.93. However, he concluded his work by indicating that cultural differences must be taken into consideration while interpreting BDI-11 item scores (Seppanen et al., 2022). Sanchez Villena et al., (2022) study used the BDI-11 to analyse the dimensionality and validate gender invariance and normative data in the Peruvian population (Sanchez Villena et al., 2022) with results indicating that the two dimensional and unifactor models had a good fit indices while bifactor and second models had convergence challenges. The unifactor model was chosen due the theory that was sound and coherent, Beck, (1961) points out that the tool should be interpreted globally using the 21 items and not separating the dimensions while calculating scores.

Psychometric Properties in Original Form

The original form of the Becks Depression inventory (Beck et al., 1961) had the following results for internal consistency at (α =.73-.92) which was evaluated as strong and test reliability test at .93 indicating it was tool which could be reliably used to assess depression symptoms. Test Retest at ($r \approx 0.60$-0.70) For content validity, the instrument was able to distinguish between psychiatric patients who exhibited and those who did not exhibit depression. the items were found measure relevant constructs within the domains of cognitive, affective and somatic with depression severity ($r \approx 0.60$-0.70) (Beck et al., 1961) indicating the ability to measure depression thus providing meaningful and accurate study results (Beck et al., 1961). The construct validity of the BDI were found to differentiate different populations, measuring psychiatric and non psychiatric patients hence the clear ability to measure the intensity of depression as a construct. For the test to pass the concurrent measure a strong score measure provides evidence that a new test is able to correctly measure a construct comparing it to the standard. Hence using the Hamiliton Rating scale, there was a strong correlation recognized to measure depression severity.

Criterion validity had a strong correlation with real world outcomes for example in the clinical settings indicating it had the capacity to reflect depressive severity (Beck et al., 1961; Makhubele et, al., 2016). The factor structure indicated two factors which include cognitive affective symptoms and somatic performance symptoms which described the relationship between the observed and unobserved constructs measured. (Beck et al., 1961). The tool was able to accurately measure depression for patients with cancer which led to the contribution of somatic items (Almeida, 2023).

However, there was a major shift in order to improve the tool from the original measures. The psychometric properties of the BDI-IA (Beck et al., 1988) indicated a shift in the internal consistence to range between (.85-.88) in both samples including clinical and non clinical population samples (Beck et al., 1988). This was an improvement from the original BDI (.73-.92) (Beck et al, 1961). According to (do Nascimento et al., 2023) the studies use the two-factor component of (cognitive- affective vs somatic/vegetative symptoms) and the factors can vary by population and the translation used for the Test-retest reliability the scores were 0.73 to 0.90 over periods ranging from one week to several months this being very dependant on the sample (Beck et al., 1988) this being a good indicator that the temporal stability of the tool for measuring depressing is high. The construct validity did not change remaining at cognitive affective and somatic performance (Beck et al., 1988; Almeida et al., 2023) using the Hamilton rating scale Hamilton (1960) and Zung Self-Rating Depression Scale Zhung (1965) to establish correlation relationship. The criterion diagnostic accuracy indicated optimal cutoffs for different types of samples. In the medical setting – oncology AUC⁼ 0,87 and cut-offs at 14-18 for specifically 14, 87% sensitivity and 73% specificity (Almeida, 2022). In order to ascertain the right clinical cut-off, context and population must be taken into consideration. The cross-cultural difference was observed when different countries were compared, item level differences indicating cultural considerations must be taken into consideration as the items perform differently, measurement invariance differs (Seppanen, 2022). The concurrent validity indicated that the tool had significant correlational values on clinical assessment and diagnosis depression. It was also able to distinguish between those who are struggling with depression and those who are not.

The tool was also able evaluate the level of sensitivity across the time, the clients who were receiving treatment indicating that the tool was a necessary and valuable tool for clinical monitoring of depression. The improvement that this tool contributed to the tool being easy to use so that those who are being assessed were able to understand statements more easily thus having the ability to get the correct responses, reducing the misinterpretation of the tool (Beck et al., 1987; Jeon et al., 2025). The shift across the use of BDI in 1961) to BDI-IA (1987) and eventually to the latest model of BDI-II (1996) wanted to clarify and reduce the overlap between (cognitions) which include appetite changes, fatigue and sleep that actual reflect as physical illness instead of picking depression thus confounding diagnosis (Knaster et al., 2016; Almeida et al., 2023).

The tool also improved and indicated strong ability to measure the depressive traits hence strong construct validity (Beck et al., 1996). Another researcher who validated the scale in Africa was (Makhubele, 2015), to assess whether the BDI-11 has the ability to measure constructs the same way in South Africa. To measure whether there were variations in the tool in relation to gender and culture (Makhubele, 2015). The participants

were from two universities with diverse characteristics and the three-factor model performed well indicating it measured the same contrast measured across other regions.

Validation in African Context

Studies conducted in sub–Saharan Africa

BDI-11 has been translated into different languages. The tool has been validated in Rwanda by researchers and as having good psychometric properties, used for assessing depression severity (Biracyaza et al., 2021; Mutabaruka et al., 2012).

Cultural Adaptations and Findings

The adaptation of the tool included the use of simple local language, used among 425 cancer patients (Biracyaza et al., 2021). This tool was accessed for the validity component (Biracyaza et al., 2021) the translated version had results which indicated high correlation between the subscales and had a high construct validity. Bartlett's test (Bartlett, 1950) indicated specificity at significant levels. The tool was reviewed for sensitivity/specificity at 0.805 able to distinguish between depressed and non depressed states (Biracyaza et al., 2021).

Post-Pandemic Psychometric Research in Africa

Post pandemic research has been carried out by different researchers in Africa. According to (Odero et al, 2023), in her study into Psychometric evaluation of PHQ-9 and Gad-7 among Community Health Volunteers and nurses, she indicated that the key determinants for a functioning health system are the Human Resources for Health (HRM) and equally important for improving individual and the population health. With the challenges of health workforce workers are continually exposed to increased stressors not limited to long working hours, high workload and poor working environments. The COVID-19 pandemic placed additional stressors on the health workers including possible infection and transmission to patients and family members, anxiety among many others (Lusambili, et al., 2023). The results indicated internal consistency of PHQ-9 and Gad with alpha and omega values exhibited above 0.7 across study samples (Odero et al, 2023). Spearman correlation for convergent and divergent validity had correlation coefficient value <0.3, 0.3 to 0.5, and above 0.5, indicating weak, moderate and strong correlation respectively. Kaiser-Meyer-Olkin (KMO) revealed an estimate of above 0.7 which was acceptable. The Confirmatory factor Analysis was used to test fit indices and Root Mean Error of Approximation (RMSEA) <0.08 acceptable fit and <0.05 good fit, Standardized Root Mean Square Residual (SRMR)<0.06, and Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) > 0.95 indicating excellent fit. The finding reveals that PHQ-9 and GAD-7 are reliable and valid tools for assessing depression and generalized anxiety, with measurement invariance across Swahili vs English versions among CHVs and nurses/midwives. This then becomes a precedent for linguistically and culturally validated mental health tools in Kenya pre and post pandemic.

In a research carried out by (Mbithi, et al., 2023) to assess the mental health and psychological well being of Kenyan adolescents, the study used PHQ-9, GAD-7, WHO-5 and Pandemic Anxiety Scale (PAS), the results reported as follows Cronbach's alpha = 0.83 for PHQ-9, and 0.81 for GAD-7. This indicated that standardized mental health tools were used and had good psychometric properties in the Kenyan Context post-COVID onset, and used on a younger population who were adolescents. In his work (Kufe et al., 2023) a study carried out in the South Africa among Health Care worker, he found the four-factor model (social dysfunction, self efficacy, capable anxiety/depression), reliability of Cronbach's alpha and McDonalds omega ~ 0.85. The study highlights the use of multi-dimensional approach in scale validation and reliability in assessing groups of professionals during the pandemic in the high-stress contexts.

Too et al. (2025) employed GAD-7, cross-language, cross culture validation during/and after COVID context and found that GAD-7 exhibited one-factor-structure in both languages of Runyoro and Luganda indicating goo internal consistency omega ~ 0.85 and correlations with PHQ. This is important as it shows how GAD-7 had good psychometric properties and successful cross cultural language validation in another African content during the pandemic era. (Gyimah, Leveana, et al, 2024)Gyimah et al. (2024), has reviewed how PHQ-9 and other tools

such as Edinburgh Postnatal Depression Scale, a 10-question self report scale have been used in low-resource settings within the African context after COVID-19 and indicates psychometric properties as being robust

The work of psychometric researcher done in Africa contributes to the growing body of knowledge post- Covid strengthening the case of culturally validated tolls used for mental Health assessment.

Validation in Kenya

Evidence from Kenyan Studies

The need for a locally validate tool adapted into the Kenyan context Abubakar, et al., evaluated the (BDI) informed Abubakar et al. (2016) on the need to research and give feedback on the tool, hence the need to examine its factorial structure. For the process to be successful, there was need to carry out in depth interviews in the first phase, (n=29) adults in attendance with the aim of the population understanding depression and this would help identify which areas of the tool needed adaptation. In phase 2 which was the next level the tool BD1-11 was administered through random selection with 221 participants being picked to give room for psychometric properties evaluation. In the third part, the study BDI-11 discriminative components were evaluated against a population of (n=29) randomly chosen Abubakar, et al., (2016) with (n=77). The re was a significant intersection with the interview results leading to significant. This then helped with the tool being modified and administered to 221 adults randomly

Abubakar, et al., (2016) confirmed the reliability and cultural relevance of the BDI-II when he translated it into the Swahili version and equally validated it. In his study he discovered the it has strong internal consistency of (α=.89). Using the two-factor structure model of the dimensions cognitive-affective and somatic symptoms. The use of culturally sensitive adaptation happened as it would not align with local realities. Once the tool was adapted it indicated good construct validity as the relationship with other constructs was acceptable.

Reliability Coefficient

Validity tests and any other adaptations (translations or cultural modifications) with the three-factor model factor giving better fit (Abubakar et al., 2016). The instrument administered had 21items was administered to n=421 adult cancer patients at Butaro Ambulatory Cancer Centre to adults aged 18 years and above. The aim was to evaluate reliability, factor structure and confirmatory analysis, the results indicated α 0.904; item correlations of between 0.342 to 0.699 signifying individual items having overall moderate to medium relationship with the scale Biracyanza et al., 2021). Validity and Factor Structure Kaiser-Maiyer-Olkin (KMO) = 0.916 correlation matrix different from identity; Confirmatory factor Analysis (CFA) of two factor model: $\chi 2 = 2{,}115.397 (p < 0.001)$ and the three-factor model at $\chi 2 = 1.699.921$ ($p< 0.001$) this model fitted better that the two-factor model (Biracyaza et al., 2021). Area Under the Curve (AUC) by using BDI-11: 0.805. Different cutoffs were reported as follows, mild depression ≥ 14; ≥ 20 moderate and ≥ 29 for major depression disorder. The corresponding sensitivity/specificity values were as follows mid cuttof: AUC = 0.781, moderate cutoff: AUC = 0.754 (Biracyaza et al., 2021). In conclusion (Biracyaza et al., 2021) emphasized the need for further validation before generalization to other populations. The instrument was used less in the medical space with the translated version being Kinyarwanda (Mutabaruka et al., 2012). The instrument was used to quantify depression symptoms severity bt reults were discussed in terms of Post Traumatic Stress Disorder (PTSD), grief and trauma exposure indicated good construct validity as the relationship with other constructs was acceptable.

Critique of Application in Kenyan Context

The tool has its strengths because of the local adaptation and validation, (Abubakar et al., 2016) especially in the low literacy populations reporting good internal consistency (α= .89) (Abubakar, et al., 2016). Qualitative interviews were used for the adaptation process hence improving the cultural relevance. Researchers in the Kenyan context have used the tool locally especially indicating feasibility. The tool hence becomes useful and relevant for severity and screening. The limitations require literacy especially being self report scale with the Likert-type options. This becomes a challenge as the tool requires a longer process which is a two-stage process

response format. Practitioners must therefore use the tool with thoughtfulness and be very consistent and transparent about the procedures for adaptations for administration use and interpretation of results.

Cultural appropriateness

The BDI-11 was adopted in the local dialect hence was able to be used for low literacy levels. Idoms used such as ''Kuchoka moyo'' (Abubarker., et al., 2016) were adapted and one, two and three factors were at the acceptable range.

Language Translations

The study does not explicitly indicate backtranslation although followed culturally sensitive processes from concept validation to pilot testing ensuring it met the standards to enable him adapt the tool to the local context (Abubakar, et al., 2016). The tool gave reliable and valid results hence can be used in the Kenyan context as a screening tool. However, there was recommendation to recalibrate the cut off points with grounding the tool to take into consideration c cultural context and cultural sensitivity.

Clinical utility

The cut off scores for the western context and population did not apply to the local Kenyan context. The scores at ''≥3 or ≥4'' Abubakar et al., (2016) was not entirely recognized.

Accessibility

Abubakar et al. (2016) adjusted the protocol administration to be used with the low literacy levels population, thus changes were made just beyond the normal simple translation. (Abubakar et al., 2016). In areas like the coastal region where literacy levels were variable, the tool was administered to adults and adolescents. It was translated into Swahili language through qualitative interviews and was considered adequate to measure depression among the Kenyan population. Kariuki et al. (2022) in his thesis has successfully used the tool and administered it within the clinical setting specifically hospital to measure depression outcomes. This indicates practicality and availability especially when permission is sought. To use the tool permission is sought and purchased though Pearson indicating that the cost may be a prohibitive factor in low-resource backgrounds hence accessibility becomes a challenge. According to Abubakar et al. (2016), because the tool has been locally validated the Swahili adapted version therefore becomes a viable option when assessing for severity index.

Ethical Concerns

The BDI-11 was originally adapted for the western population and their concepts of depression. Some of the key values that align with spirituality and interpretations or social contexts may not align with BDI-11. This therefore risks misinterpreting the local population experiences hence pathologizing what should not be. However cultural adaptation process (Abubakar et al., 2016) gives value and relevance for the tool to be used in the local context taking into consideration the participants realities. While adapting the tool (Abubakar et al., 2016) paid key consideration to ethical concerns of informed consent (APA, 2021), to enable participants to understand the reason for the assessment and why their data was important and how it would be used. The participants also had an opportunity to withdraw from the study. It is not explained how power dynamics were managed.

The chances of evoking distress while using the tool was evident, however no resources were available to explain how this ethical concern was managed. At an individual level, as a clinician I have not come across the adapted tool for my clinical practice hence concerned about equity and access hence these tools should only be used with competent trained practitioners who are able to administer and interpret the scores appropriately in line with the guidelines of the tool, over or under diagnosis will be unacceptable as it leads to harming the client.

Critique of the Application in the Kenyan Context

The development of the tool in Kenya took into consideration the gap that existed of the tool being of the western origin and did not take into consideration the Kenyan context, hence the validation of the tool was to measure

accurately, reliably and culturally appropriate depressive symptoms. The tool although accepted globally exhibited the need to validate and check its cultural relevance and psychometric properties validation within the local population and setting. There was also need to establish local norms for the Kenyan population and improve diagnostic accuracy in supporting early detection and treatment of depression across the clinical settings, community programs and schools. Abubakar et al. (2016) only sampled adult caregivers within the HIV clinical context giving room for lack of representation of diverse populations. In doing so, there was limited local validation, it was carried out with a specific location of Kilifi County among low literacy level adult population (Abubakar et al., 2016).

Mental health issues are slowly but surely becoming a challenge and the need to improve mental research in Kenya becomes a necessity. This allows researchers to collect standardized comparable data in Kenya with the aim of assessing comprehension, relevance and clarity. Hence the overall objective of validation of the BDI-11 tool in Kenya was to ensure it meets the needs of measurement of the local population including its appropriateness within the Kenyan culture, assess psychometric properties for soundness and clinical effectiveness in measuring the symptoms of depression accurately. Due to the use of this tool to only one population, specifically adolescents within the clinical sample, there is limited generalizability indicating that it does not represent the national population taking into account the diverse linguistic setting and socioeconomic status (Abubakar et al., 2017). The self reported tool depends on accurate response to the items, underreporting or over-reporting poses a challenges hence clinical interviews need to taken into consideration. The cut-off scores at minimal 0-13, mild at 14-19 were established in the western populations. There were no reports of cutoff points for the African context but reports internal consistency and discriminative validity (Abubakar et al., 2016). In the African context expression of somatic symptoms is through body symptoms rather than the western expression of depression which is a psychological language. The western context primarily focuses on western expression (Abubakar et al., 2017). Respondents who largely reside in low literacy may not accurately understand the tool thus reducing accuracy and hence adaptation to low literacy groups is of significant importance. There is therefore need to use language that is simplified specifically use of locally understood terms and items written in plain language to be understood in populations with low literacy levels. Mwangi et al. (2020), indicates that BDI-11 validation remains limited with fewer regional studies hence this poses a gap. The translation of the idioms may not fully capture the somatic expression thus there could be likelihood of misinterpretation leading to translation gaps (Abubakar, 2016).

RECOMMENDATIONS FOR PSYCHOMETRICS IN KENYA

Develop indigenous scales based on the Kenyan culture, this involves the use of assessments used in the Kenyan context is one way towards giving emphasis to the Kenyan context. Language customised to the local dialect ensures that semantics are captured well so as not lose the meaning of the word. The local dialect will capture ways of expressing emotions such that even the individuals who don't understand the English language will be able to understand the expression of emotions hence cultural adaptations. There is therefore need for cultural adaptation of the instrument ensuring somatic expressions for depression are captured (body pains, fatigue, headaches) and not only cognitive or psychological symptoms. The tool will capture and reflect social norms and values which play a key role to the locals. To strengthen local validation, studies of imported tools involve the development and validation of local tools while ensuing accepted standards or protocols for outcome validation following administration and getting a patient report (Sousa, et al., 2024). The Patient-report outcome measure (PROM) (Mokkink, 2024) allows the patient to directly assess and interpret the scores at a personal level, through the use of self-administered questionnaires thus assessing whether the right tool has been chosen (Mokkink, 2024). Translate and adopt tools into Kiswahili and local languages with rigorous psychometric testings is an important step towards instrument efficacy. The Integration of psychometrics into counselling psychology training curricula will go along way in supporting the use of tools appropriately, according to (Rwatlal, 2022) there is need to embed and prioritize psychometrics into the education curriculum to accommodate the 21st centaury challenges hence this needs to be given a priority, this specifically related to the South African context. In his conclusion, he cited the use of informed interventions and strengthening trainees' identity in addressing client social responsiveness (Rwatlal, 2022). Alipanga et al. (2022) in his research underscored the importance of expanding training to ensure safe delivery and use of psychological interventions in low- and medium-income countries. The use standardized tools requires training hence the need for platforms

that will facilitate competency-based trainings are much needed in Kenya. There is also need to establish regulatory frameworks through associations in Kenya and other relevant professional bodies. These associations will serve as gate keepers by providing peer review processes, code of ethics and the most important continuous professional development. The law bodies would form the authority as oversight bodies and the need to include psychometric assessments is key. These will licence counsellors and provide a regulatory framework for those professionals who will administer the assessments to the clients.

REFERENCES

1. Abubakar, A., Van de Vijver, F. J. R., Fischer, R., Hassan A. S., Gona J. K., Dzombo, J. T., & Newton, C. R. (2016). Adaptation and Latent Structure of the Swahili Version of Beck Depression Inventory-II in Low Literacy Population in the Context of HIV. PLoS, 11(8), e0151030.
2. Afonso, A., Portugal, A. C. D. A.. Landeira-Landeira-Fernandez, J., Bullon, F. F., Santos, E. J. R. D., Vilhena, J. D., & Anunciacao, L. (2020). Depression and Anxiety Symptoms in a Representative Sample of Undergraduate Students in Spain, Portugal, and Brazil. Psicologia: Teoria e Pesquisa, 36, e36412.
3. American Psychiatric Association. (2022). Diagnostic and Statistical Mannual of Mental Disorders (5 th ed., text rev.). American Psychiatric Association.
4. American Psychological Association. (2020). Publication of the Manual of- the American Psychological Association (7th ed.). American Psychological Association.
5. Angachi, M. O., Njiru, L., & Muturi, W. (2022). Prevalence of Anxiety and Depression among Cervical Cancer Patients in Referral Hospitals in Western Kenya. African Journal of Clinical Psychology, 01.
6. Anthony, F. Jorm & Roger T. Mulder. (2022). National Characteristics Associated with Prevalence of Depression and Anxiety Symptoms: A Cross-Sectional Ecology Study. Global Mental Health, 9, 65-71.
7. Barkham, M., Lutz, W., Lambert M. J., & Saxon, D. (2023). Routine Otcome Monitoring and Clinical Feedback in Psychological Therapy: A Review. Psychotherapy Research, 33 (1), 1-15.
8. Bartlett. M. S. (1950). Tests of Significance in factor Analysis. British Journal of Statistical Psychology, 3(2), 77-85.
9. Beck, A, T., Steer, R. A., & Garbin, M. G. (1988). Psychometric Properties of the Beck Depression Inventory: Twnty-five Years of Evaluation. Clinical Psychology Review, 8(1), 77-100.
10. Beck, A. T. (1967). Depression: Clinical Experiment and Theoretical Aspects. New York: Harper & Row.
11. Beck, A. T.,. (1961). Cognitive Therapy and Emotional Disorders. Penguin.
12. Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erabaugh, J. (1961). An inventory Measuring Depression.
13. Biracyaza, E., Habimana, S., & Rusengamihigo D. (2021). Psychometric Properties of the Beck Depression Inventory (BD1-11) in Cancer Patients from Butaro Ambulatory Cancer Center, Rwanda. Psychology Research and Behaviour Management, 14, 665-674.
14. Costa, A., Von Humboltdt S. (2021). Psychometrics, In: Gu., Dupre, M. E (eds). Encyclopedia of Gerontology and Population Aging. Springer, Cham, https://doi.org/10.1007/978-3-030-22009-9_86.
15. do Nascimento, R. L. F., Fajardo-Bullon, F., Santos, E., Landeira-Fernandez, J., & Anuncicao, L. (2023). Psychometric Properites and Cross Cultural Invariance of Beck Depression Inventory -11 and Becks Anxiety Inventory amnong Representative Sample of Spanish, Portuguese, and Brazillian Undergraduate Students. International Journal of Environmental Research and Public Health, 20(11), 6009.
16. Galton, F. (1883). Inquiries into Human Faculty and its Development. Macmillan.
17. Gori, A., Topino, E., Di Fabio, A. (2022). Assessment of Professional Self-Efficay in Psychological Interventions and Psychotherapy: A Psychometric Investigation. Journal of Clinical Psychology, 78(10), 1659-1674.
18. Gyimah, Leveana, et al. (2024). Tools for Screening Maternal Mental Health Conditions in Primary Care Settings in SuSaharan Africa: Systematic Review. Frontiers in Public Health, 12, 1321680.
19. Hamilton, M. (1960). A rating Scale for Depression. Journal of Neurology, Neurosurgery, and Psychiatry, 23(1), 56-62.
20. Hasan, M. T., & Khan, M. H. A. (2025). Psychometric Properties of Bangla Version of Becks Depression Inventory-11 (BDI-11) with Unoversity Students in Bangladesh: A cross Sectional Study. Health Science Reports, 8(1) e70280.

21. Htut, K. Z., Hlaing, P. H., Myint, A. M., & Htut, K. M. M. (2022). Factor Structure and Diagnostic Efficiency of Myanmar Version of Beck Depression Inventory-II among Substance Users. Addictive Behaviours Reports, 16, 100445.
22. Jeon, M., et al. (2025). Development and Validation of a Digital Depression Assessment Scale. Journal of Medical Internet Research, 27(1), e70689.
23. Kenya Counselling and Psychological Association. n(.d). (October 6, 2025 from https://kcpa.or.ke/about-us/). About Us. Retrieved .
24. Kenya National Bureau of Statistics. (2022. KNBS). Kenya Demographic Health Survey.
25. Knaster, P., Estlander, A, M., Karlsson, H., Kaprio, J., & Kalso, E. (2016). Daignosing Depression in Chronic Patients: DSM-IV Major Depressive Disorder vs Beck Depression Inventory BDI. PLoS ONE, 11(3), E0151982.
26. Lusambili, A., Wisofschi, S., Shumba, C., Obure, J., Mulama, K., Nyaga, L., Wade, T. j., & Temmerman, M. (2020). Health Care Workers' Perspectives of the Influences of Disrespectful Maternity Care in Rural Kenya. International Journal of Environmental Research and Public Health, 17(21), 8218.
27. Lutz, W., Rubel, J., Schwartz, B., Schilling, V., & Deis*enhofer, A. - K. (2022). Towards personalised models of psychometry outcome: A systemic Review and Synthesis. World Psychiatry, 21 (1), 18-40.
28. MacCowan, S., Shaw, S. C. .K., Doherty, M., Grosjean, B., Blank, P., & Kinnear, M. (2022). A Full CIRCLE: Inclusion of Autistic Doctors in the Royal College of Psychiatrist' Values and Equality Action Plan. The British Journal of Psychiatry (2022), 221 (1), 371 - 373.
29. Makhubela, MS., Mashegoane S. (2015). Validation of the Beck Depression Inventory -II in South Africa: Factoral Validity and Longitudinal Measurement Invariance in University Students. South African Journal of Psychology, 46(2), 203-217.
30. Makhubele, M. S. (2015 [Doctoral Dissertation, University of Limpopo] http://hdl.handle.net/10386/1195). Validation of the BDI-11 in South Africa.
31. Mbithi, G., Mabrouk, A., Odhiambo, R., Namuguzi, M., Dzombo J., & Abubakar, A. (2023). Mental Health and Psychological Well-being of Kenyan Adolescents from Nairobi and Coast Regions in the Context of COVID-19. Child and Adolescent Psychiatry and Mental Health, 17, 63.
32. Mbithi, G., Mabrouk, A., Sarki, A, et al. (2023). Mental Health and Psychological Well-Being of Kenyan Adolescents fro Nairobi and Coast Region in the Context of COVID-19. Chils Psychiatry Ment Health, 16, 63.
33. Ministry of Health. (2015). Kenyan Mental Health Policy 2015-2030. Govenrnment of Kenya Retrieved from https://www.health.go.ke.
34. Mokkink, L.B., Elsman, E. B., & Terewee, C. B. (2024). COSMIN Guidelines for Systematic Reviews of Patient-Reported Otcome Measures Version 2.0. Quality of Life Res, 33: 2929-2939.
35. Murphy, M., Wango, G., Kinanda, J, J., O'Sullivan, D., Blasi, Z .D., & Moret-Tatay, C. (2022). A Report on the Translation into Kiswahili and Subsequent Validation, of Three Positive Psychology Scales. EAS Journal of Psychology and Behavioural Sciences, 4(5) 124-133.
36. Mutabaruka, J., Sejourne N., Bui E., Birmes P, Chabrol H. (2012). Traumatic Grief and Traumatic Stress in Survivors 12 years 12 years after the genocide in Rwanda. Stress Heal, 28(8): 289-296.
37. Naja, S., Al-Kubaisi N., Chehab M., AL-Dahshan a., Abuhashem N., & Bougmiza, M. (2019). Psychometric Properties of Arabic Version of EPDS and BDI-11 as a Screening Tool for Antenatal Depression: Evidence form Quater. BMJ Open, 13(9), 9.
38. Nteere, J. G & Smith, A. (2016). Depressive Symptoms Among Emerging Adults Attending A General Outpatient Department in Kenya. African Journal of Nursing and Midwifery, 18(1) 157-169.
39. Nzangi, A., Munene, A., Chang'orok, S., & Mucherah, W. (2022). Prevalence of Depression among Adolescents in Selected Public Schools In Makueni County, Kenya. African Journal of Clinical Psychology, 4(3).Retrieved from https://repository.daystar.ac.ke/handle/123456789/5196.
40. Odero, S. A., Mwangi, P., Odhiambo, R., Nzioka, B., Shumba, C., Ndirangu-Mugo, E., Abubakar, A. (2023). Psychometric Evaluation of PHQ-9 and GAD-7 among Community Halth Volunteers and Nurses/Midwives in Kenya following a Nation-Wide Telephone Survey. Frontiers in Psychiatry, 14.
41. Rawatlal, K. V. (2022). Priorities in the 21st century Career Counselling: Implications for Counselling Psychology Training. African Journal of Career Counselling, 4(1), a59.

42. Seppanen, M., Lankila, T., Auvinen, J., Miettunen, J., Korpelainen, R., & Timonen, M. (2022). Cross-Cultural Comparison of Depressive Symptoms on the Beck Depression Inventory-11 Across Six Population Samples. BJPsych Open, 8(2), e46.
43. Silvia A., Bernardo Barahona C., Jose, O., Daniel Rodrigues da Sila., Joaquim Alves da Siva., Telmo Morinho B., Jamie G., Albino, J., & Oliveira-Maira. (2023). Criterion & Construct Validity of the Beck Depression Inventory (BDI-11) to measure Depression in Patients with Cancer- The Contribution of Somatic Items. International Journal of Clinical and Health Psychology, 23(1).
44. Smith, A., & Author, B. (2022). Cognitive Triad and Depressive Symptoms in Adolescence: A meta-analytic community analysis. Journal of Adolescent Psychology, 24, 100-120.
45. Sousa H., Ribeiro, O., Aleixo, B., Christensen, A. J., Figueiredo, D. (2024). Evaluating Psychometric Properties of Patient-Reported Outcome Measure in Assessing Symptoms in Haemodialysis: A Systematic Review Using COSMIN Guidelines. Journal of Pain and Symtom Management, 67(1).
46. Spearman, C. (1904). "Geneal Intelligence, " Objectivity Determined and Measured. The American Journal of Psychology, 15(2), 201-293.
47. Terman, L. M. (1916). The Measurement of Intelligence. Riverside Press.
48. Watson, L. K., Kaiser B. N., Ayuku D., Puffer, E. S., . (202). Validating Mental Health Assessment in Kenya Using Innovative Gold Standard. Int J Psychol, 55(3) 425-434.
49. Weschler, D. (1939). The Measurment of Adult Intelligence. William & Wilkins.
50. Wisjen, L. D., Borsboom, D., & Alexandrova, A. (2022). Values in Psychometrics, . Pesrpectives on Psychological Science, 17(4), 1045-1059.
51. Wolfgang L., Brian, S., Jamie, D. (2022). Measurement-Based and Informed Psychological Theory. Annual Review of Clinical Psychology, 18.
52. Zuffery, S., Kramer, U., Thoma, N., & Others. (2022). Using Case Formulation for prediction of the Therapeutic Allinace in Treatment for Boderline Personality Disorder. Journal of Clinical Psychology.
53. Zung, W. W. K. (1965). A Self Reporting Depression Scale. Archives of General Psychiatry, 12(1), 63-70.