
Assessing Depression and Suicidality in Asian-Americans

12

Zornitsa Kalibatseva, Ivan H.C. Wu,
and Frederick T.L. Leong

Major depression is one of the highest contributors to disability and the global burden of disease worldwide surpassing Alzheimer's and cancer (World Health Organization, 2008). Suicide is the 7th leading cause of death among men and the 15th leading cause of death among women in the USA (Heron, 2011). The serious consequences associated with depression and suicidality require increasing our understanding and familiarity with assessment tools that facilitate detection and inform appropriate intervention. The need for adequate assessment and intervention for depression and suicidality is particularly relevant for racial and ethnic minority populations in the USA given the existing mental health disparities (US Department of Health and Human Services, 2001).

The lifetime prevalence of depression among Asians and Asian-Americans varies (Kalibatseva & Leong, 2011). For example, the Chinese American

Psychiatric Epidemiological Study (CAPES) indicated a 6.9 % lifetime prevalence rate of major depressive disorder among Chinese Americans (Takeuchi et al., 1998). Similarly, 9.1 % of Asian-Americans in the National Latino and Asian-American Study (Takeuchi, Hong, Gile, & Alegría, 2007) endorsed any affective disorder in comparison to 17.9 % of non-Latino Whites, 13.5 % of Hispanics, and 10.8 % of non-Hispanic Blacks in the National Comorbidity Study-Replication (NCS-R; Breslau et al., 2006). Although Asian-Americans reported significantly lower rates of depression than European Americans, their rates seem to be higher than their overseas Asian counterparts (Chang, 2002).

While depression is found cross-culturally, the symptoms of major depression in the DSM-V (American Psychiatric Association, 2013) may not consistently and accurately capture experiences of depression across different racial/ethnic groups in the US (Kalibatseva & Leong, 2011; Ryder & Chentsova-Dutton, 2012). An important cultural consideration in the assessment of depression among Asian-Americans is the tendency to report somatic symptoms, whereas European Americans are more attuned to affective symptoms (Ryder et al., 2008). Additionally, symptoms may be presented differently as factor analytic studies have demonstrated support for fewer factors than the original solutions among ethnic minorities (e.g., Kim, DeCoster, Huang, & Chiriboga, 2011; Kuo, 1984).

Research on the assessment of depression among Asians and Asian-Americans primarily

Z. Kalibatseva, M.A. (✉)
Department of Psychology, Michigan State
University, 316 Physics Road, Rm 127B,
East Lansing, MI 48824, USA
e-mail: kalibats@msu.edu

I.H.C. Wu, M.A.
Department of Psychology, Michigan State
University, 316 Physics Road, East Lansing,
MI 48824, USA
e-mail: wuivan@msu.edu

F.T.L. Leong, Ph.D.
Department of Psychology, Michigan State
University, 316 Physics Road, Rm 136, East Lansing,
MI 48824, USA
e-mail: fleong@msu.edu

uses self-report measures and structured or semi-structured clinical interviews (Kalibatseva & Leong, 2011). Leong, Okazaki, and Tak (2003) provided a detailed review of the literature examining self-report measures of depression in East Asia. While the authors found evidence for adequate psychometric properties of the translated Western depression questionnaires, they also questioned the practice of applying current Western ethnocentric conceptualizations of depression on other racial and ethnic groups as it may result in omissions of culture-specific expression of distress (Kalibatseva & Leong, 2011).

The goal of this chapter is to identify and review the most frequently used assessment measures of depression and suicidality among Asian-Americans. When research with Asian-Americans was not available, we reviewed research with Asian samples, which may be particularly relevant for first generation Asian-Americans. Along with providing information about the reliability, validity, and utility of each measure, we discuss their strengths and weaknesses and provide norms and cut-off scores when available. The measures that we identified based on a thorough review of the depression and suicide research literature with Asian-Americans include the Center for Epidemiological Studies—Depression, Beck Depression Inventory, Geriatric Depression Scale, Hamilton Rating Scale of Depression, Patient Health Questionnaire-9, and Zung Self-Rating Depression Scale (Table 12.1).

Assessment of Depression

Center for Epidemiological Studies: Depression

The Center for Epidemiological Studies—Depression scale (CES-D; Radloff, 1977) is one of the most widely used depression questionnaires in the US. It measures the frequency of 20 symptoms of depression over the past week using a four-point Likert scale from 0 (*rarely or none of the time*) to 3 (*most or all of the time*). The questionnaire was designed to measure depressive symptoms in the general population and validated

with three White/European American samples (Radloff, 1977). Four factors emerged in these analyses, interpreted as depressed affect, positive affect, somatic symptoms, and interpersonal problems. An evaluation of the sensitivity and specificity of the CES-D suggested a cutoff of 16 to identify individuals at risk for clinical depression, although several authors recommended more stringent score cutoffs above 20 (McDowell, 2006). However, the majority of these studies recruited primarily White/European American samples and these cutoffs may differ across racial and ethnic groups.

The CES-D has been included in multiple studies with Asian and Asian-American samples although its reliability, validity, and diagnostic utility are still under research scrutiny with different Asian ethnic subgroups. One of the first studies to examine the prevalence of depression among Asian-Americans used the CES-D (Kuo, 1984). In a mixed sample of 484 Asian-Americans, the combined CES-D mean score ($M=9.38$, $SD=8.07$) was slightly higher than the mean of previously examined White/European samples (M ranges between 7.96 and 9.25). A principal component factor analysis revealed between two and three factors for each of the Asian subgroups, which differed from the four-factor solution Radloff (1977) proposed. Next, we review the CES-D research with Chinese, Japanese, Korean, and Southeast Asian-Americans to examine its validity, reliability, utility, strengths, and limitations.

Research with Chinese Samples

Ying (1988) administered the CES-D to a San Francisco community sample of 360 Chinese Americans over the phone. The measure was translated into Mandarin and Cantonese Chinese and administered in the language preferred by each participant. The mean CES-D score was 11.55 ($SD=8.23$), which was significantly higher than the Chinese American sample mean ($M=6.93$) in Kuo (1984) and the White/European sample means ($M=7.53$ – 8.58) in Radloff (1977). Ying (1988) found moderate reliability with an alpha coefficient of 0.77. Roughly one-fifth of the sample in Ying's study (24.2 %) and Kuo's study (19.1 %) scored above the suggested cutoff of 16.

Table 12.1 At-a-glance summary table

Assessment name	Condition assessed	Recommendation(s) and relevant research findings	Research conducted with:	Available in:
Beck Depression Inventory	Depression	Adequate psychometric properties	Chinese immigrants, inpatients, outpatients, and students, Asian-American college students, Asian Indian immigrants, Filipino community workers, Hmong refugees and outpatients, Japanese students, South Indian outpatients, and Thai community members	Chinese, Hmong, Korean, Japanese, Thai
Beck Depression Inventory II	Depression	Standardized Chinese version is available Adequate psychometric properties	Asian-American college students, Hong Kong students, Japanese patients and community members, Taiwanese students, and Thai police officers	Taiwanese, Chinese, Japanese, Thai
Center for Epidemiologic Studies—Depression	Depression	Mixed support for four-factor structure Bias of positive items Measurement equivalence studies are needed to assess the functional, conceptual, and metric equivalence of CES-D with various Asian groups	Chinese, Japanese, Korean, Filipino American, Vietnamese, Thai	Mandarin, Cantonese, Japanese, Korean, Vietnamese, Thai
College Student Reasons for Living Inventory	Suicidality		Asian-American college students	
Geriatric Depression Scale	Depression	Varying cut-off scores Use with caution in clinical settings Adequate psychometric properties	Chinese inpatients and outpatients, Singaporean outpatients, Japanese, Japanese Americans, and Koreans	Chinese, Japanese, Korean
Geriatric Suicide Ideation Scale	Suicidality	Adequate psychometric properties	Hong Kong Chinese	Chinese (Hong Kong)
Hamilton Rating Scale of Depression	Depression	Consistent five-factor structure Use with caution in clinical settings	Chinese outpatients and inpatients, and Japanese patients	Chinese, Japanese
Suicide Intent Scale	Suicidality	Adequate psychometric properties	Taiwanese	Chinese
Adult Suicide Ideation Questionnaire	Suicidality	Adequate psychometric properties, low positive predictive value (5.2 %)	Hong Kong Chinese adults	Chinese (Hong Kong)
Patient Health Questionnaire 9	Depression	High utility and adequate psychometric properties Questionable diagnostic ability	Chinese patients, Chinese American adults, Japanese outpatients, and Korean patients	Chinese, Japanese, Korean
Zung Self-Rating Depression Scale	Depression	Use with caution in clinical settings Adequate internal consistency and reliability Questionable factor structure	Chinese university students, Chinese elderly, and Japanese	Chinese, Japanese

Similar to Kuo's factor analysis, Ying did not replicate the four factors from Radloff's study. Instead, the Chinese Americans in both Kuo's and Ying's studies reported depressive symptoms that clustered in three factors: (1) depressed and somatic; (2) positive affect; and (3) interpersonal. This factor structure suggests that the Chinese Americans in both studies did not report somatic and affective symptoms separately.

The clustering of somatic and affective symptoms also occurred in a sample of Chinese American adolescents (Russell, Crocket, Shen, & Lee, 2008). This study examined cross-ethnic measurement invariance of the CES-D among Chinese, Filipino, and European American adolescents from the National Longitudinal Study of Adolescent Health (AddHealth). The results indicated that the four-factor solution did not fit well for the Chinese American sample. The alternative three-factor solution with 13 items offered in Ying (1988) provided a better fit of the data. Russell and colleagues (2008) concluded that Chinese American adolescents may experience distinct clusters of depressive symptoms than Filipino and European American youth. The findings suggest that the lack of differentiation between bodily and psychological depressive symptoms extends from first generation foreign-born Chinese Americans (Ying, 1988) to US-born Chinese American adolescents (Russell et al., 2008).

Li and Hicks (2010) examined the psychometric properties of the CES-D related to diagnostic validity, construct validity, internal reliability, and response bias in a probability sample of 167 Chinese American women. The CES-D revealed adequate reliability (Cronbach's $\alpha=0.86$) and 26 % of the sample had scores of 16 or above. The instrument revealed adequate construct validity as women with higher CES-D scores reported less social support, worse self-perceived general health, and more stressful life events. Moreover, women who met criteria for a current major depression diagnosis in an interview had higher CES-D scores than women who did not have a diagnosis (26.7 vs. 10.4, $p<0.001$). The study found that the CES-D cutoff of 16 or 17 ensured 100 % sensitivity (95 % CI: 44–100 %) and 76 % specificity. However, the positive

predictive value (PPV) was low (3 % for a cutoff of 16), which suggests that CES-D score above the cutoff is unreliable in confirming a diagnosis of major depression. Additionally, Li and Hicks alerted to potential response bias, such that less acculturated Chinese American women were less likely to endorse positively worded CES-D items, which artificially increased their total CES-D scores compared to more acculturated Chinese American women.

In conclusion, the CES-D demonstrated low construct validity by mixed support for its factor structure, association between low acculturation and poor validity, and cultural response bias for positive items (Kuo, 1984; Russell et al., 2008; Wong et al., 2012; Ying, 1988). While the CES-D may be a useful screening tool for nondepressed Chinese Americans, it should be followed by a diagnostic interview to identify individuals with clinical depression. The potential lack for measurement equivalence of the CES-D suggests that clinicians need to apply increased caution when administering it to Chinese Americans.

Research with Japanese Samples

Several cross-cultural studies utilized the CES-D with Japanese Americans (e.g., Kanazawa, White, & Hampson, 2007) and Japanese university students, adult workers, and adolescents (Iwata & Buka, 2002; Iwata, Roberts, & Kawakami, 1995; Iwata, Saito, & Roberts, 1994). Across all studies, Japanese and Japanese American participants had lower scores on positively worded CES-D items than European Americans. Additionally, one study found that Japanese American college students in Hawaii reported more interpersonal problems than Chinese and European American college students (Marsella, Kinzie, & Gordon, 1973). There is scarcity of studies that examine the factor structure of the CES-D with both Japanese and Japanese American samples.

Whereas we did not locate a validation study of the CES-D with Japanese Americans, the questionnaire was translated into Japanese and validated in Japan and revealed adequate clinical validity (Shima et al., 1985 as cited in Iwata et al., 1994). Additionally, the CES-D demonstrated

adequate reliability (Cronbach's α ranging from 0.76 to 0.81; Iwata et al., 1995) among Japanese samples. Kanazawa et al. (2007) also found adequate internal consistency (Cronbach's $\alpha=0.89$) for their Japanese American sample but it is important to note that these authors tested a 29-item measure as they added four items to the depressed affect factor and five items to the interpersonal problems factor. One study with Japanese workers in Japan established that 15.2 % of men and 10.6 % of women scored above 16 (Iwata, Okuyama, Kawakami, & Saito, 1989). These participants were more likely to be above age 50, never married/divorced, living alone, with bad conjugal and parental relationships, and high levels of perceived stress. We did not locate any studies that established reliable CES-D cutoffs with Japanese Americans.

Research with Korean Samples

The CES-D has been used consistently with Korean immigrants in North America. A Korean version of the CES-D was translated and validated by Noh, Avison, and Kasper (1992) with Korean immigrants in Canada. The version demonstrated content, construct, and concurrent validity and the factor structure resembled those found among American samples (Noh, Kasper, & Chen, 1998). However, several studies have suggested that Korean Americans may answer the positive items on the CES-D in a biased manner (Jang, Kim, & Chiriboga, 2005; Jang, Kwag, & Chiriboga, 2010; Oh, Koeske, & Sales, 2002). In particular, Koreans and Korean Americans were less likely to endorse the positive items, which resulted in higher overall depression scores. Since this cultural response bias has been consistently reported, some researchers have recommended calculating the total CES-D scores by excluding the positive items (Noh et al., 1998).

The reliability of the 20-item CES-D Korean version was high (Cronbach's $\alpha=0.89$) and it slightly improved in a 16-item CES-D version without the positive items ($\alpha=0.90$; Noh et al., 1998). When the positive items were reworded negatively, the reliability increased to 0.93. Additionally, the latter version demonstrated better construct validity as it correlated most highly

with another depression questionnaire and a stressful life event inventory. Thus, the revised version with 20 negatively phrased items (CESD-K-R) revealed the best reliability and validity compared to the other two versions (Noh et al., 1998).

A few studies with Korean Americans pointed out the importance of assessing acculturation when assessing depression (Jang et al., 2005; Ji & Duan, 2006; Kim, Seo, & Cain, 2006). In particular, Korean Americans who were more acculturated to American mainstream culture had a higher likelihood of endorsing the positive CES-D items similarly to European Americans, which resulted in lower overall depression scores than less acculturated Korean Americans (Jang et al., 2005). However, when acculturation was examined as bi-dimensional construct, endorsement of Korean culture was not associated with biased responding (Kim et al., 2006). We did not find any studies that examined specific CES-D norms for Korean Americans but many studies reported higher overall scores of depression among Korean Americans compared to European Americans (e.g., Jang et al., 2005; Kim et al., 2006). As previously suggested, one possible explanation for these higher scores is the differential item functioning of the positive items. Therefore, removing those items or rewording them as negative items improves the reliability and validity of the CES-D with Korean Americans. This suggestion poses questions about the utility of the CES-D.

Research with Southeast Asians Samples

There is a dearth of studies that examined the CES-D with Southeast Asian-Americans (e.g., Vietnamese, Lao, Indonesian, Malaysian, Thai, and Cambodian). We had difficulty locating psychometric studies with specific South Asian or Southeast Asian groups in the USA. The only exception was a validation study of the CES-D with Filipino American adolescents in rural and small town Hawaii (Edman et al., 1999). The reliability for this sample was high ($\alpha=0.89$) but the factor analysis results suggested a potentially different conceptualization of depression. Fifteen of the 20 items loaded on one general factor that

combined depressed affect, somatic symptoms, and interpersonal problems. The remaining four positive symptoms and the item “everything was an effort” loaded on a separate factor. Thus, although the CES-D demonstrated good reliability in this study, its factor structure with this sample of Filipino American adolescents differed from the four-factor solution.

Several studies conducted with populations from other countries included translated versions of the CES-D in Vietnam (Leggett, Zarit, Ngyuen, Hoang, & Nguyen, 2012) and Thailand (Trangkasombat & Nukhew, 1998 as cited in German et al., 2012). A cutoff of 22 was used with the Thai version to define significant level of depressive symptoms (German et al.). A cross-cultural study examined depressive symptoms using the CES-D in five Southeast Asian countries: Indonesia, North Korea, Myanmar, Sri Lanka, and Thailand (Mackinnon, McCallum, Andrews, & Anderson, 1998). There was evidence that one general depression factor fits the data better than the initially proposed four-factor solution for all samples, similar to the previously reported findings with other Asian-American samples.

Overall, the CES-D has consistently demonstrated adequate reliability and has been used successfully with community samples. Yet, researchers and clinicians need to be careful when using it with specific Asian-American populations because the proposed four-factor structure has not been replicated in a number of studies. Additionally, a substantial number of studies revealed that the positive items function differently and may need to be removed or reworded in a negative way.

Beck Depression Inventory

The Beck Depression Inventory (BDI) is a commonly used depression scale that consists of 21 symptoms and attitudes rated on intensity from 0 to 3 (Beck, Steer, & Carbin, 1988). Constructed based on a predominantly White/European American sample, the cut-off score for minimal depression is 9, mild to moderate depression is 18, moderate to severe depression is 29, and severe depression is >30 (Beck & Beamesderfer, 1974).

In the past, the English version of the BDI has been administered to Asian samples such as South Indian women in India (Chandra, Satyanarayana, & Carey, 2009), Japanese American university students ($\alpha=0.84$; Abe, 2004), Filipino Americans ($\alpha=0.87$; Napholz & Mo, 2010), and mixed samples of Asian-Americans ($\alpha=0.86$; Norasakkunkit & Kalik, 2002). The use of the BDI has expanded into studies of Asians in various countries including USA, India, China, Taiwan, Japan, Korea, and Thailand as well as various types of samples including outpatient and inpatient samples, college students, adolescents, and immigrants.

Research with Chinese Samples

In our review, we identified few standardized Chinese versions of the BDI. According to Chan (1991), the BDI has been translated into Chinese by various researchers and clinicians. These various measures appear to demonstrate adequate validity and reliability. For example, Chinese versions of the BDI have been related to post-traumatic stress disorder among Chinese workers in America (de Bocanegra & Brickman, 2004; de Bocanegra, Moskalenko, & Chan, 2005), poor self-rated health among Chinese adolescents in China (Xu et al., 2011), loneliness among Chinese university students in Taiwan (Hsu, Hailey, & Range, 1987), and hopelessness among Chinese patients (Chiles et al., 1989).

One particular version has recently undergone significant psychometric scrutiny. While Zheng, Wei, Lianggue, Guochen, and Chenggue (1988) developed a version of the BDI that yielded acceptable reliability estimates (i.e., Cronbach's $\alpha = 0.85$) and concurrent validity (i.e., moderate correlation with the Chinese Hamilton Depression Rating Scale), closer examination of the factor structure revealed three of six factors that were uninterpretable and one particularly problematic item (i.e., Item 21: libido loss). Further examining the utility of this particular version of the BDI Yeung et al. (2002) administered the questionnaire to 503 outpatient Chinese Americans (age 18+). The investigators found that the measure's optimal sensitivity and specificity was at the cut-off score of 13 and the area under the Receiver Operating Characteristic (ROC) curve was 0.94

indicating high accuracy for screening depression in their sample. Later studies conducted among this research group continued to find adequate reliability of the measure (Yeung et al., 2004, 2005; Yeung, Chang, Gresham, Nierenberg, & Fava, 2004).

Other Chinese versions of the BDI also appear to demonstrate adequate reliability. Among a group of researchers studying Chinese immigrants in the USA, reliability estimates of one translation ranged from 0.87 to 0.93 (de Bocanegra et al., 2005; de Bocanegra & Brickman, 2004; de Bocanegra, Moskalenko, & Kramer, 2006). Yet another translation revealed an overall internal reliability estimate of 0.86 and a split-half reliability of 0.77 among 2,150 Chinese adolescents in Hong Kong (Shek, 1990).

There are limited directions in guiding the interpretation of the unstandardized Chinese versions of the BDI. In fact, different cut-off scores are used depending on the study. For instance, one study of 251 Chinese inpatients in China used the cut-off scores of 13, 24, and 25 to describe normal, mild-moderate, and moderate-severe depression, respectively (Yang, Zuo, Su, & Eaton, 1987). The study found that the cut-off score of 14 in their sample best discriminated between depressed and nondepressed patients. A later study of 331 students and psychiatric inpatients in Hong Kong found the cut-off scores of 10, 19, and 30 in their Chinese version of the BDI had adequate sensitivity and specificity for determining normal, mild-moderate, and moderate-severe categories of depression, respectively, while the cut-off scores of 9, 18, and 29 in the English version of the BDI had adequate sensitivity and specificity. In both versions of the BDI, they noted that the cut-off scores for mild-moderate and moderate-severe categories had higher sensitivity and specificity than the cut-off score for the normal category (Chan, 1991).

These studies show a number of strengths and limitations among Chinese versions of the BDI. A clear strength is the strong internal consistency throughout the studies. The findings suggest that Chinese versions of the BDI, although unstandardized, demonstrate adequate reliability in screening for depression among a wide range of samples.

A notable strength of the reviewed studies is the large sample sizes. However, a number of limitations are worth noting. The lack of agreement in cut-off scores is problematic, as this limits the clinical utility of the measure. While Chan (1991) offered cut-off scores based on item analyses, further research is recommended to validate these cut-off scores. Another limitation is the lack of psychometric studies. For instance, our review did not yield any studies that conducted item response analyses or measurement equivalency for the Chinese versions of the BDI.

In summary, the studies reviewed demonstrate adequate reliability and validity for unstandardized Chinese versions of the BDI. With any unstandardized measure, however, we recommend clinicians to interpret results using Chinese versions of the BDI with caution. There lacks an agreed upon set of cut-off scores for categorizing severity, which limits the interpretability of Chinese versions of the BDI. In view of these limitations, Chinese versions of the BDI may serve as reliable supplementary assessments of depression.

Research with Hmong Samples

A Hmong version that has demonstrated adequate reliability and validity is available. The Hmong Adaptation of the Beck Depression Inventory (HABDI) made a number of changes to increase cultural sensitivity and simplicity (Mouanoutoua, Brown, Cappelletty, & Levine, 1991). The HABDI consists of 22 items rated on a three-point frequency scale rather than a severity scale found in the original BDI. An extra item was added due to translation difficulties for Item 2 (i.e., "I feel like the future is hopeless and cannot improve"). After administering the adapted version to 50 depressed and 73 nondepressed Hmong (18–66 years old) living in the US, the measure yielded a Cronbach's α of 0.93 and a 2-week test-retest reliability of 0.92. Results indicated that the depressed group scored significantly higher than the nondepressed group and a cut-off score of 46 was recommended. This cut-off score demonstrated adequate sensitivity and specificity, 94 % and 78 %, respectively.

A number of limitations are warranted in consideration of using the HABDI. First, items are

rated on a frequency scale rather than a severity scale. While the BDI has traditionally been considered a severity scale for depression, the HABDI assesses the presence or absence of symptoms limiting the ability for clinicians to determine severity. Therefore, it is recommended that the HABDI be used to supplement decisions regarding presence or absence of depression and not for severity. Second, the second item of the BDI was transformed into two items due to difficulties in translation. For instance, "I feel like the future is hopeless and cannot improve" became "I feel like the future is hopeless" and "I feel like things cannot improve." Although seen as a solution to the double-barreled item, the item analyses indicated that many endorsed the pessimism item (i.e. "I feel like the future is hopeless").

In summary, the HABDI appears to be a clinically useful tool to discriminate between depressed and nondepressed patients. However, due to the lack of empirical studies that further examine the psychometric properties of the measure, it is difficult to ascertain the degree to which this measure is reliable across the Hmong population. Thus, clinicians and researchers should interpret the findings with caution.

Research with Korean Samples

A Korean version of the BDI that has demonstrated adequate reliability is available. The BDI-K was administered to 279 Korean university students (17–37 years old) and subjected to a Rasch rating scale modeling procedure to assess the psychometric properties of the instrument (Hong & Wong, 2005). In particular, principle component analyses did not indicate any higher order factor structures despite Items 19 (i.e., weight loss) and 21 (i.e., libido loss) demonstrating low correlations with the rest of the scale. The mean score for their sample was 11.56 compared to a mean of 25.18 found in a sample of clinically depressed Koreans (Shin, Kim, & Park, 1993, as cited in Hong & Wong, 2005). The investigators further noted that the somatic items yielded greater difficulty, potentially due to the lower levels of severity in their sample. In summary, the BDI-K appears to be a reliable measure of depression among Korean university students and clinically depressed samples.

Research with Japanese Samples

A Japanese version of the BDI was translated and found to have adequate reliability among a sample of 79 female Japanese university students (Arnault, Sakamoto, & Moriwaki, 2005). The investigators found that BDI scores ranged from 0 to 39 with a mean of 12.66. Further, their scale yielded a Cronbach's α of 0.79. In a later study of 50 Japanese university women, BDI scores were significantly related to somatic distress ($r=0.57$), which explained 31 % of the variance in depression scores (Arnault, Sakamoto, & Moriwaki, 2006). In both studies, the Japanese samples had higher mean scores than the American comparative sample. Although this version consistently demonstrated adequate reliability across the two studies, further studies that examine the psychometric properties of the measure are required.

Research with Thai Samples

A Thai version of the BDI was translated and administered to a sample of community participants ($n=3,133$) in Thailand (Thavichachart et al., 2009). In an effort to examine the prevalence rates of PTSD and depression, the investigators found that their version of the BDI yielded low reliability (split-half reliability=0.74). Notably, they found that 14.3 % suffered from depression, 33.6 % from PTSD, and 11.3 % from both.

Beck Depression Inventory-II

The BDI-II (Beck, Steer, & Brown, 1996) is a revision of the earlier BDI. The BDI-II consists of 21-items that are rated based on four statements of increasing symptom severity. Distinct from the BDI, the BDI-II includes four new symptoms (i.e., agitation, worthlessness, concentration difficulty, and loss of energy) to better match the DSM-IV criteria, and a different set of cut-off scores. The cut-off scores for minimal depression is 0–13, mild depression is 14–19, moderate depression is 20–28, and severe depression is 29–63. The BDI-II is positively correlated with the Hamilton Depression Rating Scale ($r=0.71$), demonstrates high internal consistency ($\alpha=0.91$), and high 1-week test–retest reliability ($r=0.93$).

Research with Chinese Samples

Unlike the BDI, a standardized Chinese version of the BDI-II (BDI-II-C) was developed by the Chinese Behavioral Science Corporation in 2000. Previous studies find adequate internal consistency. For instance, two studies calculated Cronbach's α of 0.86 (Chang, 2005) and 0.94 (Lu, Che, Chang, & Shen, 2002) in samples of college students and outpatients, respectively. Supporting these findings, recent psychometric evaluations find that the BDI-II-C possesses strong reliability and construct validity.

Focused efforts to examine the psychometric properties support the use of the BDI-II-C in Asian samples. Using a Rasch model to examine the factorial structure of the BDI-II-C revealed two dimensions (i.e., somatic and cognitive-affective items) in a sample of 2,095 Taiwanese high school students (14–18 years old) (Wu & Chang, 2008). Similarly, a later study found that person heterogeneity did not impact the two-factor structure among a sample of 810 Taiwanese college students and found internal consistency coefficients ranging from 0.82 to 0.89 in their sample (Wu & Huang, 2010). These findings support the two-factor structure found in a sample of predominantly Caucasian university students (Storch, Roberti, & Roth, 2004). Further bolstering the psychometric properties of the BDI-II-C, a recent study of 2,922 Taiwanese adolescents (13–18.5 years old) found that overall raw scores of the two subscales were not influenced by differential item functioning across genders suggesting the use of total raw scores (Wu, 2010). Another study found that BDI-II-C items were invariant across Hong Kong and American adolescents (Byrne, Stewart, Kennard, & Lee, 2007). The investigators also found an internal consistency coefficient of 0.83 in their Hong Kong sample.

Research with Other Asian Samples

Translated into other languages, the BDI-II continued to demonstrate adequate reliability. For instance, in one study of 3,000 Thai police officers, the investigators found an internal consistency coefficient of 0.93 despite adding an extra item assessing eating habit (Chongruksa, Parinyapol,

Sawatsri, & Pansomboon, 2012). A different study translating the BDI-II into Japanese found strong internal consistency ($\alpha=0.87$) and concurrent validity ($r=0.69$ with the CES-D) among Japanese outpatients (Kojima et al., 2002). A principal component analysis also revealed a two-factor structure (i.e., somatic and cognitive-affective items) similar to previous studies in other samples and appropriate use of cut-off scores provided by Beck et al. (1996). In a later study using the same translation as Kojima et al. (2002) among a sample of outpatients, the BDI-II correlated strongly with the DSM-IV depression severity index ($r=0.77$) (Hiroe et al., 2005). The investigators recommended that a change in 5 points be considered minimal, 10–19 points be considered moderate, and >20 points be considered large clinically significant changes. Similar to the previous study, the cut-off scores provided by Beck et al. (1996) were found to be appropriate in their sample.

In summary, the reviewed studies suggest strong psychometric properties of translated versions of the BDI-II among Taiwanese, Chinese, Japanese, and Thai samples. This is consistent with one study that found adequate reliability of an English version of the BDI-II among a sample of Asian-American college students ($\alpha>0.89$) (Hambrick et al., 2010). Despite these findings, clinicians and researchers should interpret the measure with caution. Wu and Chang (2008) discovered two items that provided inadequate fit to the overall measure (i.e., Item 10—crying and Item 21—libido loss).

Geriatric Depression Scale

The Geriatric Depression Scale (GDS) is a 30-item self-report or orally administered measure that assesses behavioral and affective symptoms of depression rated on a yes/no scale. Increasing ease of administration, shorter versions of the GDS have been developed including a 15-, 12-, 10-, 5-, 4-, and 1-item version (Straus, Sherman, & Spreen, 2006). The GDS has been translated into numerous languages although its validity is questioned due to potential cultural differences in expression of depression (Jang, Small, & Haley, 2001).

Research with Chinese Samples

Existing validation studies with Chinese populations find acceptable internal reliability and validity. An early study found that the Chinese 30-item version significantly correlated with the original ($r=0.94$) and modified ($r=0.91$) 15-item version of the GDS (Liu, Lu, Yu, & Yang, 1998). The investigators also found adequate internal consistency for the original ($\alpha=0.77$) and modified form ($\alpha=0.81$). Similarly, among a sample of 461 psychiatric outpatients in Hong Kong, the Chinese version of 30-item version demonstrated adequate internal consistency ($\alpha=0.89$) and 2-week test-retest reliability ($r=0.89$) (Chan, 1996). The measure significantly correlated with the CES-D ($r=0.96$) as well as DSM-III-R diagnosis of Major Depressive Disorder ($r=0.95$).

Closer examination of the GDS, however, reveals varying cut-off scores. A study of Singaporean psychiatric outpatients suffering from dementia recommended the cut-off score of 9/10 for patients with mild dementia (sensitivity=66.7 %, specificity=95.0 %), and 11/12 for patients with severe dementia (sensitivity=40 %, specificity=100 %) (Lam et al., 2004). Another study suggested the cut-off score of 4 or above (sensitivity=84 %, specificity=85.7 %) among a community sample in Singapore (Lim et al., 2000). Yet, another study found that the cut-off score of 7/8 yielded high sensitivity (96 %) and specificity (88 %) among a community sample of depressed inpatients and outpatients (60–87 years old) in Hong Kong (Lee, Chiu, Kowk, & Leung, 1993). The varying cut-off scores are likely due to a severity effect or differences in samples. Therefore, clinicians and researchers using short form Chinese versions of the GDS are recommended to interpret cut-off scores with caution. Further, clinicians and researchers should also be wary of using the full form as a diagnostic tool (Chan, 1996).

Research with Japanese Samples

We identified one translation commonly used in the literature developed by Niino, Imaizumi, and Kawakami (1991). A recent study in Japan orally interviewed 111 elderly Japanese using the Niino et al. translation and recommended the cut-off

score of 6 based on high sensitivity (97 %) and specificity (96 %), and notable false positive (89 %) and false negative (0 %) rates (Schreiner, Hayakawa, Morimoto, & Kakuma, 2003). Supporting these findings, a different study with a community sample elderly Japanese Americans found that the cut-off score of 5 identified 20.6 % of their sample as suffering from moderate to severe depression (Mui & Shibusawa, 2003). The investigators also found a Cronbach's α of 0.86 and a split-half reliability coefficient of 0.78. The findings support the use of this measure among Japanese samples to screen for depression.

Research with Korean Samples

An early translation of the GDS short form was conducted by Jang et al. (2001) among a community sample of elderly Koreans living in metropolitan areas of Korea. The internal consistency was 0.85 and split-half reliability was 0.77. The factor structure of the Korean version of the GDS revealed the following three dimensions: (1) internal perceptions, (2) external aspects, (3) staying at home and problems with memory. While this version seemed to be reasonably useful, a later translation revealed stronger psychometric properties.

A Korean translation of the GDS developed by Bae and Cho (2004) is available and demonstrates strong psychometric properties. In this study, the full (GDS-K) and short form (SGDS-K) of the GDS was translated into Korean and administered to 154 elderly psychiatric patients in Korea. Findings of the study revealed strong correlations between the GDS-K, the SGDS-K, the Hamilton Rating Scale for Depression, and the CES-D. Further analyses indicated adequate internal reliability for the GDS-K ($\alpha=0.91$) and the SGDS-K ($\alpha=0.86$). Principal component analyses revealed the following three dimensions in the SGDS-K: (1) negative judgment about the past, present, and future, (2) lowered affect, and (3) cognitive inefficiency and lack of motivation. Overall, the researchers recommended the cut-off score of 18 for the GDS-K (sensitivity=84 %, specificity=82 %) and 10 for the SGDS-K (sensitivity=86 %, specificity=86 %).

Hamilton Rating Scale of Depression

The Hamilton Rating Scale of Depression (HRSD) is a 17-item scale that taps into various problems and symptoms of depression rated on a three- or five-point scale (Hamilton, 1960). Despite the commonly used cut-off score of 7 in clinical trials, few studies empirically examine the sensitivity and specificity of this criterion. Zimmerman, Posternak, and Chelminski (2005) found high sensitivity (97.0 %) and acceptable specificity (90.5 %) using the broad definition of DSM-IV remission in a sample of 303 psychiatric outpatients. Our review yielded few studies that sought to examine the psychometric properties of the HRSD in Asian populations.

Research with Chinese Samples

An early translation of the HRSD was conducted by Zheng, Zhao, et al. (1988) that demonstrated low reliability ($\alpha=0.71$) among a sample of outpatient and inpatients adults. Further, the scale moderately correlated with the Global Assessment Scale (GAS) ($r=-0.49$). Principal component analyses revealed the following five factors: (1) anxiety, somatization, weight-loss, (2) agitation, insight, (3) depressed mood, suicide, genital-symptoms, (4) guilt, psychomotor-retardation, and (5) sleep disturbance. A later study found that the cut-off score of 12/13 provided adequate sensitivity (88 %) and specificity (86 %) (Leung, Wing, Kwong, & Shum, 1999).

Research with Japanese Samples

A recent study conducted a cross-cultural equivalence study on the HRSD among a sample of Japanese, North American, and European patients (>18 years old) (Furukawa et al., 2005). Results indicated a somewhat consistent factor structure across the three samples. In particular, the following five factors were revealed: (1) anhedonia/retardation, (2) guilt/agitation, (3) bodily symptoms, (4) insomnia, and (5) appetite. The findings suggest an underlying factor structure across the three samples suggesting potential in cross-cultural comparisons using the HRSD.

Patient Health Questionnaire

The Patient Health Questionnaire-9 (PHQ-9) is a 9-item self-report questionnaire based on nine symptoms of depression in the DSM-IV. Each item is rated on a four-point scale indicating frequency of each symptom in the past 2 weeks. Scores ranging between 10 and 14 suggest moderate levels of depression, 15–19 suggest moderate to severe level of depression, and scores above 20 suggest severe depression (Kroenke, Spitzer, & Williams, 2001). Studies suggest that the PHQ-9 serves as an easy and short assessment for depression among Asian populations.

Research with Chinese Samples

Past studies indicate that the Chinese version of the PHQ-9 demonstrates strong psychometric properties. For instance, one study of 364 elderly Chinese patients (>60 years old) revealed high internal reliability ($\alpha=0.91$) and strong utility (7.5 min administration) (Chen et al., 2010). Further analyses indicated that the optimal cut-off of 8/9 yielded adequate sensitivity (86 %) specificity (85 %). Similarly, another study of 3,417 Chinese American adults (18–87 years old) found that 4.1 % of their sample suffered from significant depression using the cut-off score of 9/10 (Chen, Huang, Chang, & Chung, 2006). Using a higher cut-off score, a different study of 1,940 Chinese American adults (>18 years old) found high internal consistency ($\alpha=0.91$) and the cut-off score of 15 yielded moderate sensitivity (81 %) and high specificity (98 %) (Yeung et al., 2008). The authors indicated that the high cut-off score was chosen to screen for depression severity requiring psychological intervention.

Research with Japanese Samples

A study using a Japanese version of the PH-9 with a sample of 153 psychiatric outpatients identified the cut-off of 13/14 as demonstrating high sensitivity (86 %) and low specificity (67 %) for screening current major depressive episodes (Inoue et al., 2012). Further, total scores were moderately correlated with the HDRS ($r=0.55$) and the GAF ($r=-0.59$). In sum, the investigators

supported the use of the PHQ-9 for screening, but argued against the use of the scale for diagnostic purposes in their sample.

Research with Korean Samples

A Korean version of the PHQ-9 was developed by Han et al. (2008) with a sample of 1,060 elderly Korean patients (>60 years old). The investigators found that their measure demonstrated adequate internal consistency ($\alpha=0.86$) and low 3 week test–retest reliability ($r=0.79$). Further analyses indicated that the cut-off score of 4/5 yielded adequate sensitivity (80 %) and specificity (78 %) for depressive disorders. Their translation of the measure also was significantly correlated with the GDS ($r=0.74$) and the CES-D ($r=0.66$). Overall, although the Korean version of the PHQ-9 demonstrated strong criterion validity, its ability as a diagnostic tool for depression is questionable.

Zung Self-Rating Depression Scale

The Zung Self-Rating Depression Scale (SDS) is a 20-item measure that taps into characteristics of depression (Zung, Richards, & Short, 1965). Items are rated on a four-point scale based on the extent to which each item applies to the person at the time of the test. Half of the items are symptomatically positive and the other half are symptomatically negative. While the SDS has been predominantly used in White/European American samples, translations in Japanese and Chinese exist.

Research with Chinese Samples

An early study by Lee (1990) sought to examine the psychometric properties of a Chinese version of the SDS in a sample of 265 undergraduate students in Hong Kong (17–26 years old). The scale demonstrated adequate internal reliability ($\alpha=0.80$) and criterion validity evidenced by moderate correlations with the BDI ($r=0.63$), General Health Questionnaire ($r=0.58$), and the Chinese Minnesota Multiphasic Personality Inventory D-scale ($r=0.59$). Interestingly, further analyses suggested that the scale is difficult

to “fake good,” yet relatively easy to “fake bad.” To this end, the investigators cautioned the use of the SDS in clinical settings.

A later study examining the psychometric properties of the SDS among a sample of elderly Chinese in Hong Kong also found adequate reliability (Lee et al., 1994). In this sample, the SDS demonstrated strong internal consistency ($\alpha=0.91$) and split-half reliability ($r=0.89$). Furthermore, the SDS was strongly related to the GDS ($r=0.88$) and the Chinese HDRS ($r=0.86$). Overall, the cut-off score of 42/43 yielded the highest sensitivity (92.3 %) and specificity (87.5 %).

Research with Japanese Samples

A recent study of 7,136 residents in Japan (20–79 years old) were administered a Japanese version of the SDS (Chida, Okayama, Nishi, & Sakai, 2004). The investigators used the following cut-off scores: 20–39 indicated no or insignificant symptomology, 40–47 indicated mild depression, 48–55 indicated moderate depression, and >56 indicated severe depression. Results showed that 13.7 % of the sample scored within the moderate to severe range with more females (16.3 %) than males (10.6 %) in this range. Factor analyses revealed a two factor structure that only retained 12 of the 20 items in the overall sample. The researchers recommended against the use of this scale among depressed individuals reporting mainly somatic symptoms as many of these items were dropped from the factor analysis. To this end, other rating scales of depression are recommended among Japanese individuals.

Assessment of Suicidality

Suicide was the eighth leading cause of death among Asian-Americans in 2007 and the second leading cause of death for Asian-Americans in the 15–34 age group (Heron, 2011). Asian-American adults reported a lifetime prevalence of suicide ideation of 8.6 % and suicide attempts of 2.5 % (Duldulao, Takeuchi, & Hong, 2009). Choi, Rogers, and Werth (2009) provided a thorough

review of issues that may arise during the suicide risk assessment process when working with Asian-American college students. The authors recommended paying attention to issues such as self-disclosure, acculturation, intergenerational conflict, collectivistic values, the model minority myth, and perfectionism when conducting a risk assessment with Asian-American clients (Choi et al., 2009).

We found only few studies with Asian-Americans that examined the psychometric properties of suicide measures. The validity of the College Student Reasons for Living Inventory was examined with Asian-American college students (Choi & Rogers, 2010). Five of the original six factors were replicated and the reliability for each scale varied from moderate to high. The scale still explained 8 % of the variance in suicidal behavior after depression and hopelessness were already accounted for. Thus, the CSRLI seems to be a valid and reliable measure of assessing suicide risk among Asian-American college students (Choi & Rogers, 2010).

Several other measures were translated and validated with Chinese samples. The Geriatric Suicide Ideation Scale-Chinese (GSIS-C; Chou, Jun, & Chi, 2005) demonstrated adequate psychometric properties with Hong Kong older adults. Similarly, the Suicide Intent Scale had acceptable psychometric properties in a large community sample of Taiwanese respondents who committed deliberate self-harm (Gau, Chen, Lee, Chang, & Cheng, 2009). The Chinese version of the Adult Suicide Ideation Questionnaire measures the severity of suicidal ideation and also displayed appropriate psychometric properties with Hong Kong adults (Fu, Liu, & Yip, 2007). The questionnaire correctly classified 25 out of 33 people with suicidality (75 %) but had a low positive predictive value (5.2 %; cutoff score=1). The few studies that we examined reveal that a large amount of the research in this area has been conducted in the last decade and there is a need to accumulate a body of research before determining the strengths and weaknesses of specific suicidality measures.

Conclusions and Recommendations

This chapter reviewed the most commonly used assessment measures of depression and suicidality in Asian-Americans. Based on the reviewed studies, we provide a summary of our findings and suggest that future research needs to evaluate instruments' internal and external validity (Leong, 1997).

Our review is consistent with Leong et al.'s (2003) observation that a select few measures dominate the area of depression. Specifically, the BDI and the CES-D continue to receive significant psychometric scrutiny. While this increases our psychometric understanding of these measures, less attention is directed to other promising measures.

Moreover, in order to compare depression and suicidality across groups, it is necessary to demonstrate "an underlying *universal* or by searching for *equivalences*" (Berry, 1980, p. 8, italics in the original). Despite recommendations for specific procedures in the translating and adaptation of tests (International Test Commission, 2010), few studies demonstrated the measurement equivalence of each instrument. The internal validity or measurement equivalence requires that the words' meanings in another language are preserved (linguistic equivalence), the construct serve the same function across groups (functional equivalence), the conceptual frame is preserved (conceptual equivalence), and the measure's psychometric properties are transportable and generalizable to other groups (metric equivalence; Leong, Kalibatseva, & Park, 2013).

Recent research also indicated the emphasis on identifying cut-off scores and establishing the sensitivity and specificity of the used instruments. Cut-off scores and sensitivity/specificity information are especially helpful for clinicians using self-report questionnaires as supplementary diagnostic tools. Thus, instruments with adequate external validity will be able to detect correctly individuals with below clinical (i.e., specificity) and clinical levels of depression and suicidality (i.e., sensitivity). Studies, however, have largely depended on small sample sizes or

convenience samples to examine cut-off scores. Although criteria may be established for specific groups, researchers and clinicians need to use cut-off scores with caution and supplement their diagnostic decisions with data from other sources.

The increase in attention towards understanding the psychometric properties of existing measures is encouraging. This signals greater interest in validating existing measures. Yet, we noticed that it was a small pool of researchers who were conducting these validity studies. Therefore, future recommendations include establishing the measurement equivalence and specificity/sensitivity for various depression measures other than the CES-D and BDI, building on current findings and encouraging collaborations across research teams.

References

- Abe, J. A. A. (2004). Self-esteem, perception of relationships, and emotional distress: A cross-cultural study. *Personal Relationships, 11*(2), 231–247.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Arnault, D. S., Sakamoto, S., & Moriwaki, A. (2005). The association between negative self-descriptions and depressive symptomatology: Does culture make a difference? *Archives of Psychiatric Nursing, 19*(2), 93–100.
- Arnault, D. S., Sakamoto, S., & Moriwaki, A. (2006). Somatic and depressive symptoms in female Japanese and American students: A preliminary investigation. *Transcultural Psychiatry, 43*(2), 275–286.
- Bae, J. N., & Cho, M. J. (2004). Development of the Korean version of the Geriatric Depression Scale and its short form among elderly psychiatric patients. *Journal of Psychosomatic Research, 57*(3), 297.
- Beck, A., & Beamesderfer, A. (1974). Assessment of depression: The depression inventory. In P. Pichot (Ed.), *Psychological measurements in psychopharmacology: Modern problems in pharmacopsychiatry* (Vol. 7, pp. 151–169). Oxford: Karger.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
- Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*(1), 77–100.
- Berry, J. W. (1980). Introduction to methodology. In H. Triandis & J. Berry (Eds.), *Handbook of cross-cultural psychology* (Methodology, Vol. 2, pp. 1–28). Boston, MA: Allyn and Bacon.
- Breslau, J., Aguilar-Gaxiola, S., Kendler, K. S., Su, M., Williams, D., & Kessler, R. C. (2006). Specifying race-ethnic differences in risk for psychiatric disorders in a USA national sample. *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences, 36*, 57–68.
- Byrne, B. M., Stewart, S. M., Kennard, B. D., & Lee, P. W. H. (2007). The Beck Depression Inventory-II: Testing for measurement equivalence and factor mean differences across Hong Kong and American adolescents. *International Journal of Testing, 7*(3), 293–309.
- Chan, D. W. (1991). The Beck Depression Inventory: What difference does the Chinese version make? *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*(4), 616.
- Chan, A. C. (1996). Clinical validation of the Geriatric depression scale (GDS). *Journal of Aging and Health, 8*(2), 238–253.
- Chandra, P. S., Satyanarayana, V. A., & Carey, M. P. (2009). Women reporting intimate partner violence in India: Associations with PTSD and depressive symptoms. *Archives of Women's Mental Health, 12*(4), 203–209.
- Chang, D. F. (2002). Understanding the rates and distribution of mental disorders. In K. S. Kurasaki, S. Okazaki, & S. Sue (Eds.), *Asian American mental health: Assessment theories and methods* (pp. 9–27). New York, NY: Kluwer Academic/Plenum Publishers.
- Chang, H. (2005). Dimensions of the Chinese Beck Depression Inventory-II in a university sample. *Individual Differences Research, 3*(3), 193–199.
- Chen, S., Chiu, H., Xu, B., Ma, Y., Jin, T., Wu, M., & Conwell, Y. (2010). Reliability and validity of the PHQ-9 for screening late-life depression in Chinese primary care. *International Journal of Geriatric Psychiatry, 25*(11), 1127–1133.
- Chen, T. M., Huang, F. Y., Chang, C., & Chung, H. (2006). Using the PHQ-9 for depression screening and treatment monitoring for Chinese Americans in primary care. *Psychiatric Services, 57*(7), 976–981.
- Chida, F., Okayama, A., Nishi, N., & Sakai, A. (2004). Factor analysis of Zung Scale scores in a Japanese general population. *Psychiatry and Clinical Neurosciences, 58*(4), 420–426.
- Chiles, J. A., Strosahl, K. D., Ping, Z. Y., Michael, M. C., Hall, K., Jemelka, R., ... Reto, C. (1989). Depression, hopelessness, and suicidal behavior in Chinese and American psychiatric patients. *American Journal of Psychiatry, 146*(3), 339–344.
- Chinese Behavioral Sciences Society (ed.). (2000). *The Chinese version of the Beck Depression Inventory* (2nd ed., Licensed Chinese translation, The Psychological Corporation). Haccourt Brace: New York; 2000.
- Choi, J. L., & Rogers, J. R. (2010). Exploring the validity of the College Student Reasons for Living Inventory among Asian American college students. *Archives of Suicide Research, 14*, 222–235.
- Choi, J. L., Rogers, J. R., & Werth, J. L. (2009). Suicide risk assessment with Asian American college stu-

- dents: A culturally informed perspective. *The Counseling Psychologist*, 37, 186–218.
- Chongruksa, D., Parinyapol, P., Sawatsri, S., & Pansomboon, C. (2012). “Efficacy of eclectic group counseling in addressing stress among Thai police officers in terrorist situations”: Corrigendum. *Counselling Psychology Quarterly*, 25(3), 343.
- Chou, K.-L., Jun, L. W., & Chi, I. (2005). Assessing Chinese older adults’ suicidal ideation: Chinese version of the Geriatric Suicide Ideation Scale. *Aging & Mental Health*, 9, 167–171.
- de Bocanegra, H. T., & Brickman, E. (2004). Mental health impact of the World Trade Center attacks on displaced Chinese workers. *Journal of Traumatic Stress*, 17(1), 55–62.
- de Bocanegra, H. T., Moskalenko, S., & Chan, P. (2005). PTSD and depression among displaced Chinese workers after the world trade center attack: A follow-up study. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 82(3), 364–369.
- de Bocanegra, H. T., Moskalenko, S., & Kramer, E. J. (2006). PTSD, depression, prescription drug use, and health care utilization of Chinese workers affected by the WTC attacks. *Journal of Immigrant and Minority Health*, 8(3), 203–210.
- Duldulao, A. A., Takeuchi, D. T., & Hong, S. (2009). Correlates of suicidal behaviors among Asian Americans. *Archives of Suicide Research*, 13, 277–290.
- Edman, J. L., Danko, G. P., Andrade, N., McArdle, J. J., Foster, J., & Glipa, J. (1999). Factor structure of the CES-D (Center for Epidemiologic Studies Depression Scale) among Filipino-American adolescents. *Social Psychiatry and Psychiatric Epidemiology*, 34, 211–215.
- Fu, K., Liu, K. Y., & Yip, P. S. F. (2007). Predictive validity of the Chinese version of the Adult Suicidal Ideation Questionnaire: Psychometric properties and its short version. *Psychological Assessment*, 19, 422–429.
- Furukawa, T. A., Streiner, D. L., Azuma, H., Higuchi, T., Kamijima, K., Kanba, S., ... Miura, S. (2005). Cross-cultural equivalence in depression assessment: Japan–Europe–North American study. *Acta Psychiatrica Scandinavica*, 112(4), 279–285.
- German, D., Sutcliffe, C. G., Sirojrn, B., Sherman, S. G., Latkin, C. A., Aramrattana, A., et al. (2012). Unanticipated effect of a randomized peer network intervention on depressive symptoms among young metamphetamine users in Thailand. *Journal of Community Psychology*, 40, 799–813.
- Gau, S. S. F., Chen, C. H., Lee, C. T. C., Chang, J. C., & Cheng, A. T. A. (2009). Development of a Chinese version of the Suicide Intent Scale. *Suicide and Life-Threatening Behavior*, 39, 332–342.
- Hambrick, J. P., Rodebaugh, T. L., Balsis, S., Woods, C. M., Mendez, J. L., & Heimberg, R. G. (2010). Cross-ethnic measurement equivalence of measures of depression, social anxiety, and worry. *Assessment*, 17(2), 155–171.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23(1), 56.
- Han, C., Jo, S. A., Kwak, J. H., Pae, C. U., Steffens, D., Jo, I., & Park, M. H. (2008). Validation of the Patient Health Questionnaire-9 Korean version in the elderly population: The Ansan Geriatric study. *Comprehensive Psychiatry*, 49(2), 218–223.
- Heron, M. (2011). Deaths: Leading causes for 2007. *National Vital Statistics Reports*, 59, 8.
- Hong, S., & Wong, E. C. (2005). Rasch rating scale modeling of the Korean version of the beck depression inventory. *Educational and Psychological Measurement*, 65(1), 124–139.
- Hiroe, T., Kojima, M., Yamamoto, I., Nojima, S., Kinoshita, Y., Hashimoto, N., ... & Furukawa, T. A. (2005). Gradations of clinical severity and sensitivity to change assessed with the Beck Depression Inventory-II in Japanese patients with depression. *Psychiatry Research*, 135(3), 229–235.
- Hsu, L. R., Hailey, B. J., & Range, L. M. (1987). Cultural and emotional components of loneliness and depression. *The Journal of Psychology*, 121(1), 61–70.
- Inoue, T., Tanaka, T., Nakagawa, S., Nakato, Y., Kameyama, R., Boku, S., ... Koyama, T. (2012). Utility and limitations of PHQ-9 in a clinic specializing in psychiatric care. *BMC Psychiatry*, 12(1), 1–6.
- International Test Commission. (2010). *ITC guidelines for translating and adapting tests*. Accessed online on June 3, 2013, from <http://www.intestcom.org/upload/sitefiles/40.pdf>
- Iwata, N., & Buka, S. (2002). Race/ethnicity and depressive symptoms: A cross-cultural/ethnic comparison among university students in East Asia, North, and South America. *Social Science & Medicine*, 55, 2243–2252.
- Iwata, N., Okuyama, Y., Kawakami, Y., & Saito, K. (1989). Prevalence of depressive symptoms in a Japanese occupational setting: A preliminary study. *American Journal of Public Health*, 79, 1486–1489.
- Iwata, N., Roberts, C. R., & Kawakami, N. (1995). Japan–US comparison of responses to depression scale items among adult workers. *Psychiatry Research*, 58, 237–245.
- Iwata, N., Saito, K., & Roberts, R. E. (1994). Responses to a self-administered depression scale among younger adolescents in Japan. *Psychiatry Research*, 53, 275–287.
- Jang, Y., Kim, G., & Chiriboga, D. A. (2005). Acculturation and manifestation of depressive symptoms among Korean-American elder adults. *Aging & Mental Health*, 9, 500–507.
- Jang, Y., Kwag, K. H., & Chiriboga, D. A. (2010). Not saying I am happy does not mean I am not: Cultural influences on responses to positive affect items in the CES-D. *Journal of Gerontology: Psychological Sciences*, 65B(6), 684–690.
- Jang, Y., Small, B. J., & Haley, W. E. (2001). Cross-cultural comparability of the Geriatric Depression Scale: Comparison between older Koreans and older Americans. *Aging & Mental Health*, 5(1), 31–37.
- Ji, P., & Duan, C. (2006). The relationship among acculturation, acculturation stress, and depression for a

- Korean and a Korean-American sample. *Asian Journal of Counselling*, 13, 235–270.
- Kalibatseva, Z., & Leong, F. T. L. (2011). Depression among Asian Americans: Review and Recommendations. *Depression Research and Treatment*, Article ID: 320902. doi: 10.1155/2011/320902
- Kanazawa, A., White, P. M., & Hampson, S. E. (2007). Ethnic variation in depressive symptoms in a community sample in Hawaii. *Cultural Diversity and Ethnic Minority Psychology*, 13, 35–44.
- Kim, G., DeCoster, J., Huang, C., & Chiriboga, D. A. (2011). Race/ethnicity and the factor structure of the Center for Epidemiologic Studies Depression Scale: A meta-analysis. *Cultural Diversity and Ethnic Minority Psychology*, 17, 381–396.
- Kim, E., Seo, K., & Cain, K. C. (2006). Bi-dimensional acculturation and cultural response set in CES-D among Korean immigrants. *Issues in Mental Health Nursing*, 31, 576–583.
- Kojima, M., Furukawa, T. A., Takahashi, H., Kawai, M., Nagaya, T., & Tokudome, S. (2002). Cross-cultural validation of the beck depression inventory-II in Japan. *Psychiatry Research*, 110(3), 291–299.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9), 606–613.
- Kuo, W. H. (1984). Prevalence of depression among Asian-Americans. *Journal of Nervous and Mental Disease*, 172, 449–457.
- Lam, C. K., Lim, J., Ping, P. I., Low, B. L., Ng, L. L., Chiam, P. C., & Sahadevan, S. (2004). Depression in dementia: A comparative and validation study of four brief scales in the elderly Chinese. *International Journal of Geriatric Psychiatry*, 19(5), 422–428.
- Lee, H. C. B. (1990). Reliability, validity and fakability of the Zung Self-rating Depression Scale. *Bulletin of the Hong Kong Psychological Society*, 24, 5–15.
- Lee, H. C. B., Chiu, H. F. K., Kowk, W. Y., & Leung, C. M. (1993). Chinese elderly and the GDS short form: A preliminary study. *Clinical Gerontologist: The Journal of Aging and Mental Health*, 14(2), 37–42.
- Lee, H. C. B., Chiu, H. F. K., Wing, Y. K., Leung, C. M., Kwong, P. K., & Chung, D. W. S. (1994). The Zung self-rating depression scale: Screening for depression among the Hong Kong Chinese elderly. *Journal of Geriatric Psychiatry and Neurology*, 7(4), 216–220.
- Leggett, A., Zarit, S. H., Ngyuen, N. H., Hoang, C. N., & Nguyen, H. T. (2012). The influence of social factors and health on depressive symptoms and worry: A study of older Vietnamese adults. *Aging & Mental Health*, 16, 780–786.
- Leong, F. T. L. (1997). Cross-cultural career psychology: Comment on Fouad, Harmon, and Borgen (1997) and Tracey, Watanabe, and Schneider (1997). *Journal of Counseling Psychology*, 44, 355–359.
- Leong, F. T. L., Kalibatseva, Z., & Park, Y. S. (2013). Evaluating measurement equivalence in cross-cultural stress research. In P. Genkova, T. Ringeisen, & F. T. L. Leong (Eds.), *Handbook stress and culture: Intercultural and cross-cultural perspectives* (pp. 173–188). Wiesbaden, Germany: Springer Fachmedien.
- Leong, F. T. L., Okazaki, S., & Tak, J. (2003). Assessment of depression and anxiety in East Asia. *Psychological Assessment*, 15, 290–305.
- Leung, C. M., Wing, Y. K., Kwong, P. K., & Shum, A. L. K. (1999). Validation of the Chinese - Cantonese version of the Hospital Anxiety and Depression Scale and comparison with the Hamilton Rating Scale of Depression. *Acta Psychiatrica Scandinavica*, 100(6), 456–461.
- Li, Z., & Hicks, M. H. (2010). The CES-D in Chinese American women: Construct validity, diagnostic validity for major depression, and cultural response bias. *Psychiatry Research*, 175, 227–232.
- Lim, P. P. J., Ng, L. L., Chiam, P. C., Ong, P. S., Ngui, F. T., & Sahadevan, S. (2000). Validation and comparison of three brief depression scales in an elderly Chinese population. *International Journal of Geriatric Psychiatry*, 15(9), 824–830.
- Liu, C., Lu, C., Yu, S., & Yang, Y. (1998). Correlations between scores on Chinese versions of long and short forms of the geriatric depression scale among elderly Chinese. *Psychological Reports*, 82(1), 211–214.
- Lu, M. L., Che, H. H., Chang, S. W., & Shen, W. W. (2002). Reliability and validity of the Chinese version of the Beck Depression Inventory-II. *Taiwanese Journal of Psychiatry*, 16, 301–310.
- Mackinnon, A., McCallum, J., Andrews, G., & Anderson, I. (1998). The Center for Epidemiological Studies Depression Scale in older community samples in Indonesia, North Korea, Myanmar, Sri Lanka, and Thailand. *Journal of Gerontology: Psychological Sciences*, 53, 343–352.
- Marsella, A. J., Kinzie, D., & Gordon, P. (1973). Ethnic variations in the expression of depression. *Journal of Cross-Cultural Psychology*, 4, 435–458.
- McDowell, I. (2006). *Measuring health: A guide to rating scales and questionnaires* (3rd ed.). New York, NY: Oxford University Press.
- Mouanoutoua, V. L., Brown, L. G., Cappelletty, G. G., & Levine, R. V. (1991). A hmong adaptation of the beck depression inventory. *Journal of Personality Assessment*, 57(2), 309–322.
- Mui, A. C., & Shibusawa, T. (2003). Japanese American elders and the Geriatric Depression Scale. *Clinical Gerontologist*, 26(3–4), 91–104.
- Napholz, L., & Mo, W. (2010). Attribution of importance to life roles and their implications for mental health among Filipino American working women. *Health Care for Women International*, 31(2), 179–196.
- Niino, N., Imaizumi, T., & Kawakami, N. (1991). A Japanese translation of the Geriatric Depression Scale. *Clinical Gerontologist*, 10, 85–87.
- Noh, S., Avison, W. R., & Kasper, V. (1992). Depressive symptoms among Korean immigrants: Assessment of a translation of the Center for Epidemiologic Studies-Depression Scale. *Psychological Assessment*, 4, 84–91.
- Noh, S., Kasper, V., & Chen, X. (1998). Measuring depression in Korean immigrants: Assessing validity of the translated Korean version of CES-D Scale. *Cross-Cultural Research*, 32, 358–377.

- Norasakkunkit, V., & Kalik, M. S. (2002). Culture, ethnicity, and emotional distress measures. *Journal of Cross-Cultural Psychology, 33*, 56–70.
- Oh, Y., Koeske, G., & Sales, E. (2002). Acculturation, stress, and depressive symptoms among Korean immigrants in the United States. *Journal of Social Psychology, 142*, 511–516.
- Radloff, L. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.
- Russell, S. T., Crocket, L. J., Shen, Y., & Lee, S. (2008). Cross-ethnic invariance of self-esteem and depression measures for Chinese, Filipino, and European American adolescents. *Journal of Youth and Adolescence, 37*, 50–61.
- Ryder, A. G., & Chentsova-Dutton, Y. E. (2012). Depression in cultural context: “Chinese somatization”, revisited. *Psychiatric Clinics of North America, 35*, 15–36.
- Ryder, A. G., Yang, J., Zhu, X., Yao, S., Yi, J., Heine, S., & Bagby, R. M. (2008). The cultural shaping of depression: Somatic symptoms in China, psychological symptoms in North America? *Journal of Abnormal Psychology, 117*, 300–313.
- Schreiner, A. S., Hayakawa, H., Morimoto, T., & Kakuma, T. (2003). Screening for late life depression: Cut-off scores for the Geriatric Depression Scale and the Cornell Scale for Depression in Dementia among Japanese subjects. *International Journal of Geriatric Psychiatry, 18*(6), 498–505.
- Shek, D. T. L. (1990). Reliability and factorial structure of the Chinese version of the Beck Depression Inventory. *Journal of Clinical Psychology, 46*(1), 35–43.
- Shima, S., Shikano, T., Kitamura, T., & Asai, M. (1985). New self-rating scale for depression. [In Japanese] *Clinical Psychiatry, 27*, 717–723.
- Shin, M. S., Kim, Z. S., & Park, K. B. (1993). The cut-off score for the Korean version of Beck Depression Inventory. *Korean J Clin Psychol, 12*(1), 71–81.
- Storch, E. A., Roberti, J. W., & Roth, D. A. (2004). Factor structure, concurrent validity, and internal consistency of the beck depression inventory-second edition in a sample of college students. *Depression and Anxiety, 19*(3), 187–189.
- Straus, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York, NY: Oxford University Press.
- Takeuchi, D. T., Chung, R. C. Y., Lin, K. M., Shen, H., Kurasaki, K., ChinChun, C., ... Sue, S. (1998). Lifetime and twelve-month prevalence rates of major depressive episodes and dysthymia among Chinese Americans in Los Angeles. *American Journal of Psychiatry, 155*(10), 1407–1414.
- Takeuchi, D. T., Hong, S., Gile, K., & Alegría, M. (2007). Developmental contexts and mental disorders among Asian Americans. *Research in Human Development, 4*, 49–69.
- Trangkasombat, U., & Nukhew, O. (1998). Psychopathology of suicidal adolescents. *Journal of the Psychiatric Association of Thailand, 43*, 22–38.
- Thavichachart, N., Tangwongchai, S., Worakul, P., Kanchanatawan, B., Suppapatiporn, S., Pattalung, A. S., ... Chareonsook, O. (2009). Posttraumatic mental health establishment of the tsunami survivors in Thailand. *Clinical Practice and Epidemiology in Mental Health, 5*, Article 11.
- U.S. Department of Health and Human Services. (2001). *Mental health: Culture, race, and ethnicity—a supplement to mental health: A report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.
- World Health Organization (2008). The Global Burden of Disease 2004 update. Accessed online on April 28, 2013, from http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf
- Wong, R., Wu, R., Guo, C., Lam, J. K., & Snowden, L. R. (2012). Culturally sensitive depression assessment for Chinese American immigrants: Development of a comprehensive measure and a screening scale using an item response approach. *Asian American Journal of Psychology, 3*, 230–253.
- Wu, P. (2010). Differential functioning of the Chinese version of Beck Depression Inventory-II in adolescent gender groups: Use of a multiple-group mean and covariance structure model. *Social Indicators Research, 96*(3), 535–550.
- Wu, P., & Chang, L. (2008). Psychometric properties of the Chinese version of the Beck Depression Inventory-II using the Rasch model. *Measurement and Evaluation in Counseling and Development, 41*(1), 13.
- Wu, P., & Huang, T. (2010). Person heterogeneity of the BDI-II-C and its effects on dimensionality and construct validity: Using mixture item response models. *Measurement and Evaluation in Counseling and Development, 43*(3), 155–167.
- Xu, Z., Su, H., Zou, Y., Chen, J., Wu, J., & Chang, W. (2011). Self-rated health of Chinese adolescents: Distribution and its associated factors. *Scandinavian Journal of Caring Sciences, 25*(4), 780–786.
- Yang, L., Zuo, C., Su, L., & Eaton, M. T. (1987). Depression in Chinese medical inpatients. *The American Journal of Psychiatry, 144*(2), 226–228.
- Yeung, A., Chan, R., Mischoulon, D., Sonawalla, S., Wong, E., Nierenberg, A. A., & Fava, M. (2004). Prevalence of major depressive disorder among Chinese-Americans in primary care. *General Hospital Psychiatry, 26*(1), 24–30.
- Yeung, A., Chang, D., Gresham, R. L., Jr., Nierenberg, A. A., & Fava, M. (2004). Illness beliefs of depressed Chinese American patients in primary care. *The Journal of Nervous and Mental Disease, 192*(4), 324–327.
- Yeung, A., Fung, F., Yu, S. C., Vorono, S., Ly, M., Wu, S., & Fava, M. (2008). Validation of the Patient Health Questionnaire-9 for depression screening among

- Chinese Americans. *Comprehensive Psychiatry*, 49(2), 211–217.
- Yeung, A., Kung, W. W., Murakami, J. L., Mischoulon, D., Alpert, J. E., Nierenberg, A. A., & Fava, M. (2005). Outcomes of recognizing depressed Chinese American patients in primary care. *The International Journal of Psychiatry in Medicine*, 35(3), 213–224.
- Yeung, A., Neault, N., Sonawalla, S., Howarth, S., Fava, M., & Nierenberg, A. A. (2002). Screening for major depression in Asian Americans: A comparison of the Beck and the Chinese Depression Inventory. *Acta Psychiatrica Scandinavica*, 105(4), 252–257.
- Ying, Y. (1988). Depressive symptomatology among Chinese-Americans as measured by the CES-D. *Journal of Clinical Psychology*, 44, 739–746.
- Zheng, Y., Wei, L., Lianggue, G., Guochen, Z., & Chenggue, W. (1988). Applicability of the Chinese beck depression inventory. *Comprehensive Psychiatry*, 29(5), 484–489.
- Zheng, Y. P., Zhao, J. P., Phillips, M., Liu, J. B., Cai, M. F., Sun, S. Q., & Huang, M. F. (1988). Validity and reliability of the Chinese Hamilton Depression Rating Scale. *The British Journal of Psychiatry*, 152(5), 660–664.
- Zimmerman, M., Posternak, M. A., & Chelminski, I. (2005). Is the cutoff to define remission on the Hamilton Rating Scale for Depression too high? *The Journal of nervous and mental disease*, 193(3), 170–175.
- Zung, W. W., Richards, C. B., & Short, M. J. (1965). Self-rating depression scale in an outpatient clinic: further validation of the SDS. *Archives of General Psychiatry*, 13(6), 508.