The Assessment Interview: A Review of Structured and Semi-structured Clinical Interviews Available for Use Among Asian Clients

7

Diana Herva and Mary Comperini

Introduction

Of all the tools at a clinician's disposal, one of the most important is the assessment interview. Also known as the clinical interview, the assessment interview is a specialized technique that allows a clinician to glean accurate and often indispensable information regarding a client, including but not limited to the client's current condition, history, beliefs, and attitudes (Othmer & Othmer, 1994; Rogers, 2001). As such, it is important to understand the components critical for this process: building rapport, interview technique, assessment of mental status, and diagnostic needs. The purpose of the assessment interview, whether in psychological assessment or treatment, is to explain and/or classify the signs, symptoms, and behaviors that a client is exhibiting or reporting. For psychological assessments in particular, it is important for the clinician to blend the interpretation of evaluation results and the description of the symptoms (Othmer & Othmer, 1994; Rogers, 2001).

The art of the assessment interview lies in the clinician's ability to elicit essential information regarding the client. In order to do this, the clinician must have a clear idea of the evaluation's purpose and pay attention to not just content of the client's responses but also the facial expressions, tone of voice, and gestures. The assessment interview can also serve as a supportive role, which could improve rapport and help the clinician understand the client better (Othmer & Othmer, 1994; Rogers, 2001; The Department of Psychiatry Teaching Committee, 1973). Various interview styles are available and include unstructured, structured, and semi-structured interviews.

Unstructured Interviews

Traditionally, assessment interviews were unstructured. That is not to say that the clinician strides aimlessly through the assessment. Regardless of the interview style a clinician chooses to utilize, the clinician should have a clear idea of the ultimate goal (The Department of Psychiatry Teaching Committee, 1973). This understanding enables the clinician to guide the interview toward that goal. Without this guidance, the interview is unlikely to yield useful information. One of the draws of unstructured interviews is that this style is the most flexible

D. Herva, Ph.D., M.A.O.B. (✉)
Kaweah Delta Health Care District Graduate Medical Education, Visalia, CA USA

State of California Department of Developmental Services, Porterville, CA, USA
e-mail: drherva@gmail.com

M. Comperini, Psy.D.
California Department of Corrections and Rehabilitation, California Correctional Health Care Services, 1256 W Lathrop Road #351, Manteca, CA 95336, USA
e-mail: Dr.Mary@Comperini.com

L.T. Benuto et al. (eds.), *Guide to Psychological Assessment with Asians*,
DOI 10.1007/978-1-4939-0796-0_7,

due to its unrestricted nature, thereby allowing the clinician to build rapport with the client in a manner unhindered by the strict adherence to predetermined structure. However, the validity of the assessment is in question due to potential pitfalls, the most egregious of which is when a clinician prematurely terminates the interview. This can occur when the clinician comes to a conclusion regarding the client and the client's condition early on in the interview process which causes the clinician to overlook important information that could potentially disconfirm the clinician's initial opinion or that could identify additional diagnostic concerns (Othmer & Othmer, 1994; Rogers, 2001; The Department of Psychiatry Teaching Committee, 1973).

Structured Interviews

Since so many factors contribute to the validity and, thus, ultimate utility of the information obtained in an interview, structured interviews can be preferable in certain situations. Structured interviews have a predefined set of questions which provide for more reliable, quantifiable data. The structured approach has been designed rigorously to avoid biases in the line of questioning and allow for trained nonprofessionals to conduct the interview without missing important information (Rogers, 2001; Vatrapu & Perez-Quinones, n.d.). Standardized and systematic administration and scoring is clearly a major advantage of structured interviews. This comprehensive approach decreases the variability in interviewing and increases diagnostic accuracy. However, the drawback is its rigidity. Structured interviews do not allow for deviations from the language and sequence of the measurement. Therefore, a clinician cannot provide additional prompts and/or explanations without being concerned about the impact on the validity of the assessment. Even if a clinician adheres to the structure, there is still potential for validity concerns, especially if the clinician uses the structured interview in an unskilled manner. Strictly adhering to the protocol can lead to the unintended effects of missing important nonverbal information, damage of rapport due to the impersonal nature of the interview, and overwhelming the client with questioning (Rogers, 2001).

Semi-structured Interviews

Semi-structured interviews are a happy medium between the unstructured and structured interview styles. They blend the strengths of both styles: diagnostic accuracy and flexibility. Unfortunately, they also blend the weaknesses of both styles. Regardless, the use of semi-structured interviews may be more appealing when working with a minority population, especially non-English speakers as it allows much more flexibility in regards to the questions, optional prompts, and ability to explain or elaborate on particular points in the interview.

Recommendations

Regardless of the interview style that a clinician prefers, the clinician should keep in mind that establishing and maintaining rapport with the client is essential for obtaining relevant information. Additionally, the clinician should be careful not to miss important nonverbal cues, such as gestures, tone of voice, and hesitations. It could be beneficial to include screening measures in an assessment interview can help to identify individuals with severe mental illness as it could suggest the need for further assessment and treatment. Finally, the clinician should be flexible with his conclusions regarding diagnosis to avoid a confirmation bias. This can occur when the clinician simply seeks information to confirm the diagnosis rather than considering the possibility of alternative or additional diagnoses as well.

Cultural Considerations

Culturally sensitive interviews will take into consideration the client's characteristics and preferences as well as potential limitations. Skill and sensitivity are of utmost importance in this

essential process of evaluation (Othmer & Othmer, 1994; Rogers, 2001; The Department of Psychiatry Teaching Committee, 1973). When working with diverse populations, it is critical that the clinician consider the stigma attached to mental health and mental health-related services in general. Sue and Sue (1987) explained that "the amount of stigma or shame associated with emotional difficulties is probably much greater among Asian-American groups. Mental illness in a family member is considered a failure of the family system itself" (p. 480). Even among Asian individuals of later generations, there is a low mental health utilization rate, often only seeking treatment for severe symptomology (Okazaki, 1998; Sue & Sue, 1987). Therefore, the clinician should consider the purpose of the evaluation and be aware that Asian-Americans may conceptualize mental health problems in different way, primarily via somatic complaints. The somatic manifestation of symptoms complicates the diagnostic process. For example, the DSM criteria for depression requires dysphoric mood as a primary symptom, which would exclude those Asian-Americans who express depression somatically (Sue & Sue, 1987). The different manifestations have often been discussed along with the prevalence of symptoms, symptom patterns, and disorders (Okazaki, 1998).

Another significant factor that clinicians should consider is the potential for language barriers. The APA's 1990 Guidelines for Providers of Psychological Services to Ethnic, Linguistic, and Culturally Diverse Populations recommends that the clinician interact with the client in the client's preferred language. If the clinician lacks this language skill, an appropriate referral or use of an interpreter should be considered. When using an interpreter, the clinician should ensure that the interpreter is knowledgeable about the culture, has appropriate professional background, and has no dual roles with the client. It is of particular importance that the clinician adheres to the latter as much as possible as it may affect the validity of the entire evaluation.

While research has shown that interpreter services improve overall health care experiences and outcomes, clients who required the use of interpreters were more likely to have questions that they wanted to ask but did not (Green et al., 2005), which could have significant implications regarding the client's level of understanding and ability to respond fully to questions asked of them. Clients also appear to engage more with individuals of the same culture than with someone of Anglo-American backgrounds (Vatrapu & Perez-Quinones, n.d.). Green et al. (2005) suggest that "[u]se of interpreters may also compromise rapport between patients and clinicians, and their presence may inhibit patients' questions, particularly about sensitive topics such as mental health" (p. 1054). Green et al. (2005) also found that clients who rated their interpreters highly were more likely to also rate overall care highly. Thus, the role of the interpreter is crucial to a client's access to health care, including mental health. It is important to keep in mind that when using an interpreter, the assessment may take more time or there may not be as much time for the client to ask for clarification. Clinicians who utilize interpreters should be cognizant of this and allow clients an opportunity to ask questions without feeling rushed. Clinicians should also be aware that there may not be corresponding words in the client's preferred language, which could reduce the effectiveness of the interview (Sue & Sue, 1987). Therefore, it is particularly important to refrain from using jargon and to speak in short, clear sentences in order to ensure understanding (Green et al., 2005). The clinician should also question overly simplistic responses as well as responses that fail to address the intended query as these could be indications that the client lacks adequate understanding of the question or that the question was inadequately translated.

The paucity of research on the Asian-American population is a barrier to the establishment of clear guidelines for culturally competent assessment and interpretation of assessment measures that are typical in a standard assessment battery. Partially, this is due to the sociodemographic shifts and within group differences (Okazaki, 1998; Sue & Sue, 1987). As such, it is critical that the clinician be culturally sensitive which means that the clinician has enough knowledge and understanding of a client's culture to

adequately make distinctions between pathological and nonpathological symptomology (Okazaki, 1998). Despite the limited availability of structured and semi-structured interview instruments in Asian languages, they can nonetheless be useful. Brief measures, such as screening tools, appear to be well received by the Asian population (Veijola et al., 2003). In determining whether a measure would be appropriate to use with a particular population or subpopulation, however, the clinician must also consider the available psychometric properties of such measures.

Brief Review of Psychometric Properties

Psychometric properties such as reliability and validity are important in determining the usefulness of a measure. Reliability refers to the consistency of the results over time (test-retest reliability) and between different raters at the same point in time (interrater reliability). In regards to categorical constructs, such as psychiatric diagnoses as assessed on some psychological measures, is reported in terms of kappa, a statistic that corrects for chance agreement. Generally, kappa values above 0.70 are considered to reflect good agreement, kappa values between 0.50 and 0.70 are considered to be fair, and kappa values under 0.50 are considered to be poor.

Validity refers to the extent that a test measures what it claims to measure, particularly whether the test items cover what is being measured (content validity), whether the test is effective in predicting a construct (criterion-related validity), and whether the test is related to what it is meant to measure (construct validity). When psychometric properties are limited or absent, the clinician must make a decision whether to use such a measure or not by weighing the potential utility of the tool against the lack of data for its use with such a population. The validity of a diagnostic instrument is often measured by comparing agreement between diagnoses made by the instrument and diagnoses made via a "gold standard." It should be noted that the gold standard has yet to be definitively identified in regards to psychiatric diagnoses. Some studies use psychiatrist expert diagnoses while others use available comprehensive measures such as the Structured Clinical Interview for the DSM-IV Disorders (SCID) and World Health Organization (WHO) Composite International Diagnostic Interview (CIDI).

Structured and Semi-structured Assessment Measures

Given the complexities of an assessment interview, it would be beneficial to know what assessment measures are available. While there are many measures available for use with various populations, this chapter will focus on some of the more frequently used measures that are applicable to the assessment interview and is not meant to be a comprehensive or inclusive list of available measures. The included measures were selected based on their availability, accessibility, and utility for the purposes outlined in this chapter. Considering the diversity of the Asian-American culture and languages, these measures are not available in all languages and psychometric properties such as norms, reliability, and validity data are limited. Where available, these properties are discussed below.

In the event that such languages and norms are unavailable but a clinician chooses to utilize such measure, the clinician should be aware of the possible limitation of using such a measure on the desired population and have knowledge of the measure's reference population (APA, 2002). Any test results should be interpreted with caution while considering the client's cultural and linguistic characteristics. It is the clinician's responsibility to exercise critical judgment when deciding to use assessment measures not currently normed or linguistically available for a specific population.

At-a-glance summary table

Assessment name	Disorder assessed	Recommendation(s) and/or relevant research findings	May be used with	Available in
Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)	Adjustment Disorders, Affective Disorders, Anxiety Disorders, Eating Disorders, Substance Use Disorders	Good for diagnosing schizophrenia and mood disorders in Chinese inpatients. Generally good reliability.	Chinese Americans, Korean Americans	Mandarin, Korean
Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II)	DSM-IV Personality Disorders	Reasonable validity with Cantonese-speaking and Japanese individuals. Excellent reliability in Thai population.	Japanese Americans, Korean Americans, Chinese Americans, Thai Americans	Japanese, Korean, Mandarin, Thai
Mini-Mental State Examination (MMSE-II)	Global Cognitive Status	Generally good psychometric properties. May not be appropriate for individuals with low education and/or literacy levels.	Chinese Americans, Indian Americans, Japanese, Americans, Korean Americans, Thai Americans, Vietnamese Americans	Bengali, Chinese (Cantonese and Mandarin), Gujarati, Hindi, Indian English, Japanese, Kannada, Korean, Malay, Malayalam, Marathi, Punjabi, Tamil, Telugu, Thai, Urdu, Vietnamese
Diagnostic Interview Schedule (DIS)	Affective Disorders, Anxiety Disorders, Antisocial Personality Disorder, Disruptive Behavior Disorders, Eating Disorders, Psychotic Disorders, Substance Use Disorders	Moderate concordance rates, but generally reliable.	Chinese Americans	Chinese
Hopkins Symptom Checklist-25 (HSCL-25)	Affective Disorders, Anxiety Disorders, Unspecified Distress	Use as a screening tool in traumatic contexts to measure unspecified distress. Well received by Southeast Asian population. Cutoff point = 1.75.	Southeast Asians (Cambodian, Hmong, Laotian, Vietnamese), Japanese Americans, Tibetans	Cambodian, Hmong, Japanese, Laotian, Tibetan, Vietnamese
Composite International Diagnostic Interview (CIDI)	Affective Disorders, Anxiety Disorders, Dissociative Disorders, Eating Disorders, Psychotic Disorders, Substance Use Disorders	Limited data on psychometric properties for translated versions. Identifies cases with fewer indictors.	Chinese Americans, Japanese Americans, ndian Americans	Chinese, Japanese, some Indian languages
Mini International Neuropsychiatric Interview (MINI)	Affective Disorders, Anxiety Disorders, Psychotic Disorders, Substance Use Disorders	Limited data on psychometric properties for translated versions.	Chinese Americans, Japanese Americans, Korean Americans, Malaysian Americans, Taiwanese Americans, Filipino Americans, Thai Americans	Cantonese, Japanese, Korean, Malay, Mandarin, Tagalog, Thai

Structured Clinical Interview for the DSM-IV Axis I and II Disorders (SCID-I and SCID-II)

The Structured Clinical Interview for the DSM-IV Axis I and II Disorders (SCID-I and SCID-II) are semi-structured interviews designed to diagnose major DSM-IV disorders (Biometrics Research Department, n.d.). The SCID-I has two main versions: a Clinical Version (SCID-CV) and several research versions. The SCID-CV was designed for efficient use in clinical settings while the research versions are more comprehensive. The primary difference in the SCID-CV and the SCID-I Research Version is the level of detail that is covered. Additionally, the SCID-I Research Version incorporates the following diagnoses, which are not included in the SCID-CV: Acute Stress Disorder, Minor Depressive Disorder, Mixed Anxiety Depressive Disorder, and Binge Eating Disorder. The SCID-I Research Versions include an extensive patient edition (SCID-I/P) for use with psychiatric patients and a shorter nonpatient edition (SCID-I/NP), abridged from the SCID-I/P, for use with subjects not identified as psychiatric patients. The SCID-II was designed to assess the 11 major Axis II personality disorders in the DSM-IV-TR, including Personality Disorder NOS, Depressive Personality Disorder, and Passive-Aggressive Personality Disorder (Biometrics Research Department, n.d.; Summerfeldt, Kloosterman, & Antony, 2010).

Validity, Reliability, and Utility

Biometrics Research Department (n.d.) reports reliability data from various sources indicating generally fair to good reliability and validity of the SCID-I and SCID-II. Interrater reliability kappa coefficients ranged from 0.48 to 0.98 and internal consistency coefficients were fair to good (0.71–0.94) (Biometrics Research Department, n.d.; Maffei et al., 1997). Available psychometric properties for versions adapted for use with Asian populations appear to demonstrate similar reliability and validity. The *Chinese-Bilingual SCID-I/P* (CB-SCID-I/P) demonstrated an overall kappa for interrater reliability of 0.84; specifically, the kappa for bipolar affective disorder was 0.84, mood disorder was 0.76, and schizophrenia was 0.75. The overall kappa value was 0.77 and the percentage of agreement was 89.6 %. The CB-SCID-I/P was determined to be a reliable instrument for diagnosing schizophrenia and mood disorders for inpatients (So et al., 2003). The *Borderline Personality Disorder Subscale* (*Chinese Version*) *of the SCID-II* was demonstrated to have internal consistency of 0.82 (Cronbach's alpha) with an agreement kappa of 0.82, sensitivity of 0.92, and specificity of 0.94. The SCID-II Borderline Personality Disorder Subscale (Chinese Version) was determined to have reasonable validity for use with Cantonese-speaking individuals (Wong & Chow, 2011). The *Japanese version* of the SCID-II personality questionnaire (SCID-II-PQ) has been demonstrated to have moderately good test-retest reliability, which increased to being very good following the SCID-II interview. The SCID-II was determined to be useful and reliable for the Japanese population. The overall kappa for the SCID-II-PQ was 0.56 and the overall kappa for the SCID-II was reported at 0.87 (Osone & Takahashi, 2003). The *Thai version* of the SCID-II has been shown to have good to excellent reliability with an interrater reliability for diagnoses ranging between 0.70 and 0.90 (Wongpakaran et al., 2012).

Although it is unclear if there is a Khmer translation available, Hinton, Ba, Peou, and Um (2000) reported that the Cambodian Panic Disorder Survey (CPDS) is a culturally valid adaptation of the SCID-panic module. Research provides for the prevalence and subtypes of panic disorders for this population (Hinton et al., 2000).

Interpretation

Normative data is not readily available. However, research conducted with translated versions of the SCID has shown it is generally a reliable measure (Osone & Takahashi, 2003; So et al., 2003; Wong & Chow, 2011; Wongpakaran et al., 2012). Research has shown that the SCID may be more conservative in identifying cases than screening measures although the SCID and screening measures tended to agree in regards to noncases (Silov et al., 2007).

Strengths and Limitations

As a comprehensive measure, the SCID instruments have clear advantages of being able to assess a broad scope of diagnoses in a hierarchical approach. These measures have good diagnostic reliability. Conversely, limitations in these measures include the length of assessment due to its comprehensive nature. Additionally, the focus on diagnosis can lead to insufficient attention to key symptoms. The SCID provides clinical and subclinical gradations, but may not be useful in evaluating changes in symptom severity.

Special Considerations

A SCID version has been developed for use in children (KID-SCID) with generally good reliability (Biometrics Research Department, n.d.; Summerfeldt et al., 2010). Unfortunately, research with this measure and its availability in other languages is limited.

Languages

The SCID-I is available in Korean and Mandarin while the SCID-II is available in Japanese, Korean, Mandarin, and Thai.

Specific Recommendations

The clinician should keep in mind cultural factors that may affect diagnoses, particularly as the SCID measures are based on the DSM-IV criteria. Additionally, with the recent release of the DSM-V, the clinician should be aware of changes to the diagnostic criteria that are not yet reflected in these measures.

Mini-Mental State Examination (MMSE)

No assessment interview is complete without a basic mental status examination. The MMSE is a structured measure that is commonly used to evaluate global cognitive status. The most current version, MMSE 2nd Edition (MMSE-2), is reported to be equivalent to the original MMSE and includes brief, standard, and expanded forms. The MMSE-2 expanded the MMSE to also be used with milder forms of cognitive impairment.

Validity, Reliability, and Utility

The *Japanese version* has been shown to have good reliability (Cronbach's coefficients exceeding 0.70) and three main factors (immediate memory, orientation and delayed recall, and working memory) that explain 44.6 % of the variance (Shigemori, Ohgi, Okuyama, Shimura, & Schneider, 2010). The Japanese version was determined to accurately reflect the cognitive profile of older Japanese adults and the question about orientation to time may be useful in the simplest assessment to identify cognitive dysfunction (Ideno, Takayama, Hayashi, Takagi, & Sugai, 2011; Shigemori et al., 2010). There are three *Korean versions*: full Korean MMSE (K-MMSE), MMSE optimized for screening dementia (MMSE-DS), and a short version of the MMSE-DS (SMMSE-DS) (Kim et al., 2010). The MMSE-DS was reportedly optimized for screening dementia. The SMMSE-DS was constructed from the MMSE-DS based on the diagnostic accuracy of items for dementia. The internal consistency (Cronbach's coefficient alpha = 0.826), interrater reliability (0.968), and test-retest reliability (0.825) were good. The MMSE-DS is reported to have significant correlation with the Clinical Dementia Rating and the K-MMSE. Research indicates that, as compared to the K-MMSE, the SMMSE-DS had increased sensitivity and specificity for dementia (Kim et al., 2010).

Interpretation

Although specific psychometric properties are not readily available on the *Cantonese version of the MMSE* (C-MMSE), research has been conducted on a telephone version of the C-MMSE (T-CMMSE) which suggests excellent inter- and intra-rater reliability (Wong & Fong, 2009). The agreement between the face-to-face MMSE and T-CMMSE was reported with a kappa of 0.60–0.80 for orientation, registration, and recall items. Wong and Fong (2009) suggested a cutoff score of 16 for the T-CMMSE for discrimination between clients with and without dementia.

In a study of Korean neurosurgical patients, the mean score on the *Korean version* (K-MMSE) was 22.3 with 62.1 % of scores

below 23 (Kim et al., 2010). Based on the administration of the Mini-Mental State Examination from the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment Packet (CERAD-K) Neuropsychological Assessment Packet (MMSE-KC), Kim et al. (2012) reported demographics-adjusted norms for the MMSE in elderly Koreans, suggesting that "clinicians need to take age, education, and gender into account for an accurate interpretation of the MMSE-KC total score" (p. 5). Please refer to Kim et al. (2012) for specific norms.

The mean score on the *Thai version* (MMSE-T) in general non-dementia Thai subjects was reported at 27.2 and ranged between 17 and 30; 7.44 % of the cases were reported to have MMSE-T scores above 23 (Wongchaisuwan, Sithinamsuwan, Udommongkol, & Wongmek, 2005). However, research suggests that the MMSE-T may not be appropriate as a screen for cognitive impairment in clients of lower literacy and education levels (Jitapunkul & Lailert, 1997; Wongchaisuwan et al., 2005).

PAR Publishers (2012) caution that users of non-English language versions of the MMSE-2 should base the clinical interpretation of the MMSE-2 scores upon locally collected standardization data and/or clinical patient data. The publishers indicate that they do not endorse the use of the U.S. population-based MMSE-2 norms for interpretation in the MMSE-2 scores of patients who do not match the demographic characteristics of the MMSE-2 standardization sample. "Clinical use of the raw sore cutoff ranges for the MMSE-2 should be based on the scientific literature on the MMSE/MMSE-2 in the native language in which it is being used" (PAR, 2012).

Strengths and Limitations

The obvious strength of this measure is its brief nature, which is likely part of the reason that it is so commonly used. However, the MMSE is a screening tool and does not provide diagnoses. Research has also shown that it may not be appropriate for all populations, especially in those with low education levels.

Special Considerations

The MMSE is generally used with adults and older adults as an assessment of global cognitive status. It is important to remember that this is a screening tool that should not be used independently for diagnostic purposes.

Languages

The MMSE-2 is available in simplified Chinese and Hindi. The original MMSE also has authorized translations in Bengali, Cantonese Chinese (for Hong Kong), Chinese, Chinese (for Hong Kong, Malaysia, Singapore, and Taiwan), Gujarati, Hindi, Indian English, Japanese, Kannada, Korean, Malay, Malayalam, Mandarin (also for China, Singapore, and Taiwan), Marathi, Punjabi, Tamil, Telugu, Thai, Urdu, and Vietnamese.

Specific Recommendations

If the MMSE is unavailable in the language necessary, other options may be to utilize select subtests from available translated versions of intelligence assessment measures such as the Wechsler Adult Intelligence Scales and Stanford-Binet (Zuckerman, 2005). Such subtests can include Information, Arithmetic, Comprehension, Similarities, and Digit Span. This method allows the clinician to assess a client's mental status while allowing for precise scoring and interpretation. However, the clinician should be cognizant of the fact that norming and validity may still be impacted due to the unstandardized manner of administering such intelligence tests (Zuckerman, 2005).

Diagnostic Interview Schedule (DIS)

The DIS is an extensive structured diagnostic interview that was designed to assess for current and lifetime diagnoses, which are organized into 19 diagnostic modules. Information obtained in the DIS includes onset, duration, and recency of symptoms. The DIS was originally designed for the purpose of determining the prevalence and incidence of certain diagnoses in the United Stated. Since its inception, the DIS has demonstrated its cross-cultural applications and utility in both

clinical and research arenas. The structured nature of the DIS allows both professional and nonprofessional interviewers to administer it (Rogers, 2001).

Validity, Reliability, and Utility

The original Chinese version was reported to have a moderate concordance rate (kappa = 0.54). Studies generally indicate that the Chinese version is a reliable tool (Hwu & Chang, 1986; Hwu, Yeh, Chang, & Yeh, 1986).

Interpretation

Research with the Chinese version have found lower prevalence rates for certain disorders (Hwu et al., 1986; Hwu & Compton, 1994), which is reasonable considering that some psychiatric conditions are naturally less prevalent in Asian-Americans than in other populations.

Strengths and Limitations

Over the years, the DIS had been used as the gold standard due to its extensive and structured nature. It is easily administered by professionals and nonprofessionals alike. Additionally, the DIS has a self-administered version in two different formats (computer-assisted and paper-and-pencil), further simplifying the assessment process. However, like many assessment measures, the DIS is lacking in normative and validation data for translated versions, which can be significant considering cultural differences. Regardless, research has shown that the DIS is a good cross-cultural measure that has been widely used. In fact, the DIS so comprehensive and widely used that it was expanded into the WHO-CIDI by incorporating the International Classification of Diseases (ICD) diagnostic criteria.

Special Considerations

The DIS is meant for adults ages 18 and over. However, there is a Chinese version of the Diagnostic Interview Schedule for Children-Version 4 (DISC-IV). This version can be used on children and adolescents. Kappa coefficients ranged from 0.37 to 0.84 with three diagnoses falling under 0.50 (anxiety disorder, ADHD, and ODD). All three low kappa levels were found when the youth was the informant. Despite this, research indicates that the Chinese version had comparable test-retest reliability with the original English version (Ho et al., 2005).

Languages

The DIS is available in Chinese and Korean.

Specific Recommendations

As with any measure without specific norms or validation data, the DIS should be used with caution and the clinician should keep in mind the cultural differences that may affect diagnoses.

Hopkins Symptom Checklist-25 (HSCL-25)

The Hopkins Symptom Checklist (HSCL) is a semi-structured self-report screening instrument that was originally designed to measure change in a refugee client's clinical status, primarily unspecific distress. The HSCL-25 was developed for use in primary care settings from the 90-item Symptom Checklist (SCL-90), which was designed to assess patterns of current psychological symptoms (Mollica, Wyshak, de Marneffe, Khuon, & Lavelle, 1987; Ventevogel et al., 2007). It includes 10 items from the HSCL-58 anxiety cluster, 13 items from the depression cluster, and 2 additional somatic symptoms. The items are rated on a 4-point scale from "not at all" to "extremely." Depending on the purpose for using this measure, the entire measure or its subsections can be used separately. The HSCL-25 can also be administered by trained nonprofessionals.

The HSCL-25 was developed for use with the Southeast Asian population due to the lack of mental health care for refugees in the United States. In the 1980s, three versions were originally developed: Cambodian, Laotian, and Vietnamese. In 1995, a Hmong version became available (Mollica et al., 1987; Mounoutoua & Brown, 1995). The Japanese version was reportedly developed for use in the Kobe earthquake study (Mollica et al., 1987).

Validity, Reliability, and Utility

The original *Cambodian*, *Laotian*, *and Vietnamese versions* were validated together due to a small sample size for each language group alone (Mollica et al., 1987). Mollica et al. (1987) found that the HSCL-25 was sufficiently sensitive and specific for the presence of depression. The HSCL-25 was also found to be correlated with clients' self-assessments of clinical improvement. In regards to reliability, the "test-retest coefficients for the three language groups combined were 0.89 for the total score and 0.82 for anxiety and depression; the results for each language group separately were comparable. The interrater reliability for the total, anxiety, and depression scores for each of the three language groups was higher than 0.98" (Mollica et al., 1987, p. 499). Convergent reliability has been measured at 0.73–0.88 with measures of depression. Mollica et al. (1987) reported that these results were comparable to previous findings on the HSCL-25.

The *Hmong version* is reported to have "internal consistency of 0.97 and had a split-half coefficient of 0.92 and test-retest reliability of 0.90… The Hmong version of the HSCL-25 provided a sensitivity of 100 %, specificity of 78 % [resulting in a false positive rate of 22 %], and overall accuracy of 89 % [for distinguishing clinical and nonclinical groups], demonstrating that it is a useful screening tool for assessing general distress and anxiety problems in Hmong people" (abstract, p. 1). The Hmong version also used a cutoff score of 1.75 (Mounoutoua & Brown, 1995). The *Tibetan version* is reported to have coefficient alphas between 0.87 and 0.89 for the anxiety subscale and 0.85–0.92 for the depression subscale (Keller et al., 2006; Lhewa, Banu, Rosenfeld, & Keller, 2007). Results indicated good classification accuracy for anxiety (0.89), depression (0.92), and PTSD (0.83) (Lhewa et al., 2007). Research has reported prevalence rates for anxiety (25–77 %), depression (11.5–57 %), and PTSD (11–23 %) in Tibetan refugees (Mills et al., 2005). The *Japanese version* (J-HSCL-25) has been reported to have test-retest reliability coefficients of 0.75 for depression and 0.78 for anxiety with a cutoff score of 1.75. The Cronbach's alphas ranged from 0.90 to 0.91 for depression and 0.84–0.86 for anxiety (Mollica, McDonald, Massagli, & Silove, 2004; Sumi & Kanda, 2002).

Interpretation

For the original *Cambodian*, *Laotian*, *and Vietnamese versions*, a cutoff point of 1.75 was selected "for its consistency with data obtained from a random general population sample in California" (Mollica et al., 1987). Though most research studies have used 1.75 as the cutoff, Mollica et al. (1987) submit that there is no psychometric evidence to suggest that 1.75 is the ideal cutoff point. Research indicates that lower cutoff scores may be more sensitive and be able to accurately classify those individuals with PTSD (Keller et al., 2006). Regardless of the ambiguity surrounding the ideal cutoff point, these translated versions were found to be comparable to other HSCL-25 versions in regards to sensitivity, specificity, and reliability (Mollica et al., 1987; Mounoutoua & Brown, 1995; Sumi & Kanda, 2002).

Though the research for the available Asian language versions used cutoff scores of 1.75, other research suggests that the cutoff scores determined in one cultural group should be reconsidered in other groups (Ichikawa, Nakahara, & Wakai, 2006). Ichikawa et al. (2006) examined the use of the predetermined cutoff score of 1.75 on Afghan refugees in Japan. The results indicated that the use of this cutoff score overestimated the magnitude of mental health problems in this population. An algorithm method could improve classification accuracy as compared to using cutoff scores; however, its clinical validity remains unknown. Ichikawa et al. (2006) surmised that the overestimation in mental health problems as measured by the HSCL-25 may be due to the differences in response to the symptoms, causing noncases to be classified as cases. Therefore, considering the within group differences in the Asian population, the clinician should be cautious about applying the same cutoff score for subpopulations not yet represented in research. Nevertheless, as previously mentioned, research conducted with all the versions discussed above used a cutoff score of 1.75.

Strengths and Limitations

The HSCL-25 is a good screening measure for use with the Southeast Asian population. This is primarily due to its clinical utility, briefness, and economical nature (Sandanger et al., 1999). This measure has been shown to be readily accepted by Southeast Asians due to its similarity to medical tests (Mollica et al., 1987). Additionally, its linguistic simplicity allows clients of various educational levels to easily comprehend the items. The 4-point likert-type scale from "not at all" to "extremely" allows for measurement of clinical change in clients. Additionally, the HSCL-25 has been shown to have a sensitivity of 48 %, specificity of 87 %, and sensitivity of comorbid psychiatric disorders of 100 % for DSM-III-R psychiatric disorders; this suggests that it is a moderate instrument for screening (Veijola et al., 2003). Although studies on DSM-IV criteria are unavailable for review, this data suggests that this measure is adequate for screening purposes, particularly for cases that are sensitive to pain, distress, and impairment as well as for symptoms related to somatic illnesses (Sandanger et al., 1999). In fact, Silov et al. (2007) found that the HSCL-25 agreed with the SCID in identifying noncases and suggested that it may be more accurate in identifying cases in highly symptomatic clinic populations.

A limitation of this measure is that it does not provide a diagnosis. On the other hand, the clinician can gain useful information in regards to symptoms of anxiety and depression, which can help determine if deeper probing is warranted. Moreover, research has shown that the HCSL-25 can be helpful in evaluating victims of trauma, particularly for Southeast Asians who have experienced serious traumas. The design of this measure allows for the assessment of trauma in a nonthreatening manner that helps clients verbalize their symptoms.

Special Considerations

The HSCL-25 has been used primarily in traumatic contexts with adults and adolescents as young at 14 years of age.

Languages

The HSCL-25 is available in the Cambodian, Hmong, Japanese, Laotian, Tibetan, and Vietnamese languages.

Specific Recommendations

As with all screening tools, it is important to keep in mind that the purpose is to determine whether more in depth evaluation of such symptoms is warranted. Therefore, despite the potential limitations of this measure, it can provide the clinician with indications of areas for further exploration.

Composite International Diagnostic Interview (CIDI)

The CIDI is a structured measure that was developed in 1990 by the World Health Organization (WHO) as an expansion of the Diagnostic Interview Schedule (DIS). Its history can be traced back to the early 1980s when a WHO task force considered combining the DIS and Present State Examination (PSE) in a concerted international effort (WHO, n.d.; Wittchen, 1994). The CIDI expanded on the DIS by incorporating the International Classification of Disease (ICD) diagnostic system. The rationale was so that cross-national comparative research could be conducted without limitation to just mental disorders. In 1998, the CIDI was further expanded to include measurements for risk factors, consequences, treatment, etc. This measure evaluates lifetime and 12-month disorders (WHO, n.d.; Wittchen, 1994).

There are two versions of the CIDI available: the computer-assisted version (CAPI) and the paper and pencil version (PAPI). Both versions are modularized so the clinician can select any or all of the sections for administration as well as selecting the percentage of subjects who will randomly enter certain sections. The CAPI offers a user-friendly interface that simplifies the process of selecting sections (so that clinicians can choose which sections to assess as well as the

percentage of subjects who will randomly enter certain sections). The PAPI formerly allowed for screening via the short form, which could be followed up by a long form. However, the WHO no longer supports the use of the CIDI short form as it has been determined to not be useful. The PAPI version also incorporates a reference card to assist the clinician in making decisions to skip or administer certain sections without having the clinician flip back and forth between pages for a client's previous responses (WHO, n.d.).

The CIDI Diagnostic Algorithms include: mood, anxiety, substance-related, and impulse control disorders. Included in mood disorders are: Major Depressive Disorder, Recurrent Brief Depression, Hypomania, and Mania. Anxiety disorders include: Agoraphobia, Generalized Anxiety Disorder, Panic Attack/Disorder, Social Phobia, Specific Phobia, Separation Anxiety Disorder, and Post Traumatic Stress Disorder (PTSD). Substance-relate disorders include: Alcohol Abuse/Dependence, Drug Abuse/Dependence, and Nicotine Dependence. Impulse Control Disorders include: Conduct Disorder, Intermittent Explosive Disorder, Anorexia Nervosa, Bulimia Nervosa, Pathological Gambling, Attention Deficit Disorder, Hyperkinetic Disorder, and Oppositional Defiant Disorder. The CIDI also include a diagnostic algorithm for Premenstrual Dysphoric Disorder (WHO, n.d.; Wittchen, 1994).

Validity, Reliability, and Utility

In general, the WHO-CIDI has been demonstrated to have "good to excellent Kappa coefficients for most diagnostic sections. In international multicenter studies as well as several smaller center studies the CIDI was judged to be acceptable for most subjects and was found to be appropriate for use in different kinds of settings and countries. There is however still a need for reliability studies in general population samples, the area the CIDI was primary [sic] intended for. Only a few selected aspects of validity have been examined so far, mostly in smaller selected clinical samples" (Wittchen, 1994, abstract). Test-retest kappa values ranged from 0.59 to 0.84 and interrater reliability kappa values ranged from 0.67 to 0.98 (Wittchen, 1994). Diagnostic concordance rates have been found to be relatively high (kappa values range between 0.73 and 0.83). However, these studies were involved small sample sizes; the CIDI lacks large-scale validation and no norms are available. Despite this, the CIDI has been widely translated and used in many countries. Part of its vast appeal is that the CIDI is based on both DSM and ICD diagnostic criteria, which allows for cross-national comparative research and utility. As such, the CIDI has been used for validation of other measures including the HSCL-25 and MINI (Lecrubier et al., 1997; Sandanger et al., 1999).

The Chinese version of the WHO-CIDI version 3.0 (CIDI-3.0) was reported to be an acceptably validated instrument for community survey on mental disorders (Huany et al., 2010). The screen section of the CIDI-3.0 had sensitivity values of 60.4–93.1 % and specificity values of 33.6–92.7 %. As for the different mental disorders, sensitivity values ranged from 33.3 to 70.3 % and specificity values ranged from 97.1 to 98.9 %. Consistency of any mental disorder was reported at 0.78 and test-retest reliability kappa values ranged from 0.74 to 1.00 (Huany et al., 2010).

Interpretation

Interpretation is simplified in that scoring algorithms for this measure are available to certified users and are computer-based, which minimizes user error. Despite the lack of large-scale validation and norms, this tool has been widely used and is determined to be useful for diagnostic purposes. The WHO has strict guidelines for the translation of this measure and translations are expected to be equivalent to the English version. Therefore, interpretation should be similar across cultures.

Strengths and Limitations

Copies of the CIDI are readily available at no charge. Formal training at an official TRC is required before the CIDI can be used. This training requires 30 h of at-home self-study followed by 3–5 days of in-person training and a certification assessment. The scoring algorithms are only available upon becoming a certified user. One of the strengths of this measure is that it can be administered by trained laypersons as well as clinicians.

Despite the comprehensiveness of this measure, one of the largest limitations is the extensive amount of time it takes to administer, especially if the entire measure is administered. As a result of the modularization, clinicians can significantly shorten the administration time if only specific sections are selected.

Special Considerations

The CIDI can be used with individuals ages 16 and older. Though there are no norms available, this measure has been widely translated and used. The WHO's strict translation requirements suggest that translated versions are equivalent to the English version. Additionally, since the CIDI is based on both DSM and ICD diagnostic criteria, this measure should, theoretically, be just as useful for translated versions.

Languages

The CIDI is available in Chinese, Japanese, and limited access in some Indian languages. However, the WHO cautions that many of the translated "versions were done for specific studies and are out of date" (WHO, n.d.).

Specific Recommendations

While this measure has been widely translated and used, the clinician should keep in mind the purpose of the evaluation. Research has indicated that the HSCL-25 and the CIDI identified the same amount of cases though the agreement of identified cases was only half (Sandanger et al., 1999). The cases identified by the HSCL-25 had significantly more illness indicators and the cases identified by just the CIDI had the least. Thus, the clinician must weigh the benefits of the measure against potential time constraints.

Mini-International Neuropsychiatric Interview (MINI)

The MINI is a short diagnostic structured interview developed by psychiatrists and clinicians in the United States and France. The primary purpose for its development was to create a short measure based on DSM. However, the current version (MINI 6.0) is based on both DSM-IV and ICD-10 criteria and screens for 17 Axis I disorders. This measure was designed in such a manner that it could be administered by nonspecialized interviewers in about 15 min. As such, this measure is widely used internationally for psychiatric evaluation and outcome tracking (Lecrubier et al., 1997; Medical Outcome Systems, n.d.; Sheehan et al., 1998).

Validity, Reliability, and Utility

The MINI has been validated against the SCID, CIDI, and expert opinions. It has been lauded as a fully validated and time-efficient measure that can be easily integrated into clinical interviews. In fact, research demonstrates that the MINI was better at diagnosing comorbid conditions. Additionally, it has been shown to be well accepted by clients (Medical Outcome Systems, n.d.; Pinninti, Madison, Musser, & Rissmiller, 2003; Sheehan et al., 1998). In general, sensitivity and specificity were considered to be "good" (sensitivity = 0.45–0.96, specificity = 0.86–1.00). Agreement between MINI diagnoses and experts was found in 85 % of cases. Test-retest reliability kappa values were between 0.35 and 1.00 with only Current Mania falling under 0.50, which is understandable considering that mania is a changing state. The interrater reliability kappa values were between 0.79 and 1.00 (Sheehan et al., 1998).

For the *Japanese version*, "[a]ll of the kappa values indicated excellent agreement (more than 0.75) except those for dysthymia (0.74) and generalized anxiety disorder (0.72)… Kappa values indicated excellent agreement (more than 0.75) for 3 of the 12 diagnoses (i.e. major depressive disorder, panic disorder, and generalized anxiety disorder), good agreement (0.60–0.74) for six diagnoses, and acceptable agreement (0.45–0.59) for two diagnoses. Only one kappa value (dysthymia) was less than 0.45" (Otsubo et al., 2005, p. 521). Otsubo et al. (2005) also found that in general, the MINI diagnoses had acceptable to excellent agreement with the SCID-P results (0.45 to more than 0.75); 3 of the 16 diagnoses were excluded because no subjects met the criteria for these diagnoses (PTSD, alcohol abuse, and drug abuse).

Additionally, there was poor concordance rate (<5 %) for the following diagnoses: mania, generalized anxiety disorder, drug dependence, and bulimia nervosa (Otsubo et al., 2005). Overall, these data appear similar to, if not better than, the validation data for the original MINI. Although data on reliability and validity are unavailable for the other translated versions of the MINI, it can be surmised that the translated versions would be equivalent to the English versions and, thus, retain similar psychometric properties.

Interpretation

Although widely used, this measure is lacking in adequate norms and validation for the translated versions. However, considering that it is based on both DSM and ICD criteria, there is a high likelihood that interpretation is similar across cultures.

Strengths and Limitations

For clinicians in private practice, this measure is available for a one-time fee and there is no per use fee. Additionally, there are paper and electronic versions available. The electronic version can be used directly from any browser. Clinicians can have clients complete the MINI at home prior to the in-office visit.

Clearly, the lack of norms and validation for translated versions of this measure are limitations that raise serious concerns about its applicability across cultures. However, as previously discussed, since it is based on both DSM and ICD diagnostic criteria, there is a high likelihood that the MINI would be just as useful diagnostically across cultures.

Special Considerations

This measure has a version (MINI-KID) developed specifically for use with children and adolescents. The MINI-KID version has a high concordance rate for mood disorders, anxiety disorders, substance use disorders, ADHD/behavioral disorders, and eating disorders (kappa values between 0.56 and 0.87) while results were more variable for psychotic disorders (kappa = 0.41). Sensitivity was between 0.61 and 1.00 and specificity was between 0.81 and 1.00. Test-retest and interrater reliability were deemed high (kappa values between 0.64 and 1.00). Overall, the MINI-KID appears to be reliable for use on children and adolescents. It should be noted that this version may not be widely available in other languages (Sheehan et al., 2010).

Languages

The MINI is available in Cantonese, Japanese, Korean, Malay, Mandarin Chinese for Taiwan, Tagalog, and Thai.

Specific Recommendations

The clinician should use caution when using this measure due to the lack of available norms and validation data. However, it could be a useful tool as it is based on both DSM and ICD criteria.

Summary/Conclusion

The assessment interview is essential to the evaluation process and allows a clinician to glean accurate and often indispensable information regarding a client, including but not limited to the client's current condition, history, beliefs, and attitudes (Othmer & Othmer, 1994; Rogers, 2001). Therefore, it is imperative that a clinician understand the components critical for this process: building rapport, interview technique, assessment of mental status, and diagnostic needs. The measure selected for use and interview style can greatly affect the validity and utility of data obtain from the interview itself. Screening measures can be used to help identify individuals with severe mental illness as it could suggest the need for further assessment and treatment while more comprehensive measures can assist with diagnostic needs.

Culturally sensitive interviews will take into consideration the client's characteristics and preferences as well as potential limitations. The paucity of research with the Asian population prevents the establishment of clear guidelines for culturally competent assessment and interpretation of assessment measures that are typical in a standard assessment battery. As such, it is critical that the clinician be culturally sensitive and have

enough knowledge and understanding of a client's culture in order to adequately make distinctions between pathological and nonpathological symptomology (Okazaki, 1998). Clinicians should consider the purpose of the evaluation and be aware that Asian-Americans may conceptualize mental health problems in different way, namely somatic complaints, which further complicate the diagnostic process. Another significant factor that clinicians should consider is the potential for language barriers and the need for interpreter service as well as the possible complications that are involved.

The Asian-American population is so diverse that it is difficult to find, translate, or validate measures for all subpopulations. Even with the few that were identified and discussed in this chapter are limited linguistically and psychometrically. Therefore, it is inevitable that clinicians will require the use of psychological measures that have not been validated on the appropriate populations. In such an event, the clinician should be aware of the possible limitations of using such measures on the desired population and have knowledge of the measure's reference population (APA, 2002). The American Psychological Association's amended 2010 Ethical Principles of Psychologists and Code of Conduct specify that "[w]hen such validity or reliability [for the test population] has not been established, psychologists describe the strengths and limitations of test results and interpretation" (p. 12). In addition to this, test results should be interpreted with caution while considering the client's cultural and linguistic characteristics. Ultimately, it is the clinician's responsibility to exercise critical judgment when deciding to use assessment measures not currently normed or linguistically available for the specific population. In making this decision, the clinician should make an effort to seek consultation and proceed in a manner that is most in line with ethical standards. It is recommended that the clinician first attempt to identify alternative measures that may be appropriate for the intended purpose before attempting to modify existing measures for use. Clinicians should be aware that any modifications to the standardized administration of a measure may affect the test results; such limitations should be discussed and incorporated into the interpretation of the results.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Association.

American Psychological Association. (2002, August). *Guidelines on multicultural education, training, research, practice, and organizational change for psychologists* [Publication]. Retrieved from http://www.apa.org/pi/oema/resources/policy/multicultural-guidelines.aspx

American Psychological Association. (2010, February). *Ethical principles of psychologists and code of conduct with the 2010 amendments*. Retrieved from http://www.apa.org/ethics/code/index.aspx

American Psychological Association, Office of Ethnic Minority Affairs. (1990, August). *Guidelines for providers of psychological services to ethnic, linguistic, and culturally diverse populations* [Publication]. Retrieved from http://www.apa.org/pi/oema/resources/policy/provider-guidelines.aspx

Biometrics Research Department. (n.d.). *Structured clinical interview for DSM disorders (SCID)*. Retrieved from http://www.scid4.org/index.html

Green, A. R., Ngo-Metzger, Q., Legedza, A. T. R., Massagli, M. P., Phillips, R. S., & Iezzoni, L. I. (2005). Interpreter services, language concordance, and health care quality experiences of Asian Americans with limited English proficiency. *Journal of General Internal Medicine, 20*, 1050–1056.

Hinton, D., Ba, P., Peou, S., & Um, K. (2000). Panic disorder among Cambodian refugees attending a psychiatric clinic. *General Hospital Psychiatry, 22*(6), 437–444.

Ho, T., Leung, P. W., Lee, C., Tang, C., Hung, S., Kwong, S., … Shaffer, D. (2005). Test-retest reliability of the Chinese version of the Diagnostic Interview Schedule for Children-Version 4 (DISC-IV). *Journal of Child Psychology and Psychiatry*, *46*(10), 1135–1138.

Huany, Y., Shou-Fu, X., Jin, L., Jun-Ting, X., Wei-Min, D., Yi, L., … Rui, C. (2010). Community-based evaluation of the reliability and validity of Chinese version of Composite International Diagnostic Interview-3.0. *Chinese Mental Health Journal, 24*(1), 21–24.

Hwu, G. H., & Chang, L. Y. (1986). Chinese diagnostic interview schedule: I. Agreement with psychiatrist's diagnosis. *Acta Psychiatrica Scandinavica, 73*(3), 225–233.

Hwu, G. H., & Compton, W. M. (1994). Comparison of major epidemiological surveys using the Diagnostic Interview Schedule. *International Review of Psychiatry, 6*(4), 309–327.

Hwu, G. H., Yeh, E. K., Chang, L. Y., & Yeh, Y. L. (1986). Chinese diagnostic interview schedule, II. A validity study on estimation of lifetime prevalence. *Acta Psychiatrica Scandinavica, 73*, 348–357.

Ichikawa, M., Nakahara, S., & Wakai, S. (2006). Cross-cultural use of the predetermined scale cutoff points in refugee mental health research. *Social Psychiatry and Psychiatric Epidemiology, 41*, 248–250.

Ideno, Y., Takayama, M., Hayashi, K., Takagi, H., & Sugai, Y. (2011). Evaluation of a Japanese version of the Mini-Mental State Examination in elderly persons. *Geriatrics & Gerontology International, 12*(2), 310–316.

Jitapunkul, S., & Lailert, C. (1997). Mini-Mental Status Examination: Is it appropriate for screening in Thai elderly? *Journal of the Medical Association of Thailand, 80*(2), 116–120.

Keller, A., Lhewa, D., Rosenfeld, B., Sachs, E., Aladjem, A., Cohen, I., … Porterfield, K. (2006). Traumatic experiences and psychological distress in an urban refugee population seeking treatment services. *The Journal of Nervous and Mental Disease, 194*(3), 188–194.

Kim, T. H., Jhoo, J. H., Park, J. H., Kim, J. L, Ryu, S. H., Moon, S. W., … Kim, K. W. (2010). Korean version of Mini Mental Status Examination for Dementia Screening and its' [sic] short form. *Korean Neuropsychiatric Association, 7*(2), 102–108.

Kim, J. L., Park, J. H., Kim, B. J., Kim, M. D., Kim, S., Chi, Y. K., … Kim, K. W. (2012). Interactive influences of demographics on the Mini-Mental State Examination (MMSE) and the demographics-adjusted norms for MMSE in elderly Koreans. *International Psychogeriatrics, 24*(4), 642–650.

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., …, Dunbar, G. C. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. *European Psychiatry, 12*(5), 224–231.

Lhewa, D., Banu, S., Rosenfeld, B., & Keller, A. (2007). Validation of a Tibetan translation of the Hopkins Symptom Checklist 25 and the Harvard Trauma Questionnaire. *Assessment, 14*(3), 223–230.

Maffei, C., Fossati, A., Agostoni, I., Barraco, A., Bagnato, M., Deborah, D., … Petrachi, M. (1997). Interrater reliability and internal consistency of the structured clinical interview for DSM-IV axis II personality disorders (SCID-II), version 2.0. *Journal of Personality Disorders, 11*(3), 279–284.

Medical Outcome Systems. (n.d.). M.I.N.I. 6.0. *Medical Outcomes Systems: A diagnosis and treatment outcome tracking solutions company*. Retrieved from https://medical-outcomes.com/index/mini

Mills, E. J., Singh, S., Holtz, T. H., Chase, R. M., Dolma, S., Santa-Barbara, J., & Orbinski, J. J. (2005). Prevalence of mental disorders and torture among Tibetan refugees: A systematic review. *BMC International Health and Human Rights, 5*, 7.

Mollica, R. F., McDonald, L. S., Massagli, M. P., & Silove, D. M. (2004). *Measuring trauma, measuring torture: Instructions and guidance on the utilization of the Harvard Program in Refugee Trauma's versions of the Hopkins Symptom Checklist-25 (HSCL-25) and The Harvard Trauma Questionnaire (HTQ)*. Cambridge, MA: Harvard Program in Refugee Trauma.

Mollica, R. F., Wyshak, G., de Marneffe, D., Khuon, F., & Lavelle, J. (1987). Indochinese versions of the Hopkins Symptom Checklist-25: A screening instrument for the psychiatric care of refugees. *American Journal of Psychiatry, 144*(4), 497–500.

Mounoutoua, V. L., & Brown, L. G. (1995). Hopkins Symptom Checklist-25, Hmong Version: A screening instrument for psychological distress. *Journal of Personality Assessment, 64*(2), 376–383.

Okazaki, S. (1998). Psychological assessment of Asian Americans: Research agenda for cultural competency. *Journal of Personality Assessment, 70*(1), 54–70.

Osone, A., & Takahashi, S. (2003). Twelve month test-retest reliability of a Japanese version of the Structured Clinical Interview for DSM-IV Personality Disorders. *Psychiatry and Clinical Neurosciences, 57*(5), 532–538.

Othmer, E., & Othmer, S. C. (1994). *The clinical interview using DSM-IV Volume 1: Fundamentals*. Washington, DC: American Psychiatric Press.

Otsubo, T., Tanaka, K., Koda, R., Shinoda, J., Sano, N., Tanaka, S., … Kamijima, K. (2005). Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. *Psychiatry and Clinical Neurosciences, 59*, 517–526.

PAR. (2012). *Mini-Mental State Examination, 2nd Edition: Simplified Chinese (MMSE-2)*. Retrieved from http://www4.parinc.com/Products/Product.aspx?ProductID=MMSE-2:CHINESE

Pinninti, N., Madison, H., Musser, E., & Rissmiller, D. (2003). Mini International Neuropsychiatric Schedule: Clinical utility and patient acceptance. *European Psychiatry, 18*(7), 361–364.

Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., & Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research, 18*(3), 169–184.

Rogers, R. (2001). *Handbook of diagnostic and structured interviewing*. New York, NY: The Guilford Press.

Sandanger, I., Moum, T., Ingebrigtsen, G., Sorensen, T., Dalgard, O. S., & Bruusgaard, D. (1999). The meaning and significance of caseness: The Hopkins Symptom Checklist-25 and the Composite International Diagnostic Interview II. *Social Psychiatry and Psychiatric Epidemiology, 34*, 53–59.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a

structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*(Suppl. 20), 22–33.

Sheehan, D. V., Sheehan, K. H., Shytle, R. D., Janavs, J., Bannon, Y., Rogers, J. E., … Wilkinson, B. (2010). Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). *The Journal of Clinical Psychiatry, 71*(3), 313–326.

Shigemori, K., Ohgi, S., Okuyama, E., Shimura, T., & Schneider, E. (2010). The factorial structure of the Mini Mental State Examination (MMSE) in Japanese dementia patients. *BMC Geriatrics, 10*(36), 1–7.

Silov, D., Manicavasagar, V., Mollica, R., Thai, M., Khiek, D., Lavelle, J., & Tor, S. (2007). Screening for depression and PTSD in a Cambodian population unaffected by war: Comparing the Hopkins Symptom Checklist and Harvard Trauma Questionnaire with the structured clinical interview. *Journal of Nervous and Mental Disease, 195*(2), 152–157.

So, E., Kam, I., Leung, C. M., Chung, D., Liu, Z., & Fong, S. (2003). The Chinese-Bilingual SCID-I/P project: Stage 1-reliability for mood disorders and schizophrenia. *Hong Kong Journal of Psychiatry, 13*(1), 7–18.

Sue, E., & Sue, S. (1987). Cultural factors in the clinical assessment of Asian Americans. *Journal of Consulting and Clinical Psychology, 55*(4), 479–487.

Sumi, K., & Kanda, K. (2002). Relationship between neurotic perfectionism, depression, anxiety, and psychosomatic symptoms: A prospective study among Japanese men. *Personality and Individual Differences, 23*, 817–826.

Summerfeldt, L. J., Kloosterman, P. H., & Antony, M. M. (2010). Structured and semistructured diagnostic interviews. In M. M. Antony & D. H. Barlow (Eds.), *Handbook of assessment and treatment planning for psychological disorders* (2nd ed., pp. 95–140). New York, NY: The Guilford Press.

The Department of Psychiatry Teaching Committee. (1973). *Notes on eliciting and recording clinical information*. Oxford: Oxford University Press.

Vatrapu, R., & Perez-Quinones, M. A. (n.d.). Culture and international usability testing: The effects of culture in structured interviews [article]. Retrieved from http://arxiv.org/ftp/cs/papers/0405/0405045.pdf

Veijola, J., Jokelainen, J., Lasky, K., Kantojarvi, L., Kokkonen, P., Jarvelin, M., & Joukamaa, M. (2003). The Hopkins Symptom Checklist-25 in screening DSM-III-R axis I disorders. *Nordic Journal of Psychiatry, 57*(2), 119–123.

Ventevogel, P., Vries, G. D., Scholte, W. F., Shinwari, N. R., Faiz, H., Nassery, R., … Olff, M. (2007). Properties of the Hopkins Symptom Checklist-25 (HSCL-25) and the Self-Reporting Questionnaire (SRQ-20) as screening instruments used in primary care in Afghanistan. *Social Psychiatry and Psychiatric Epidemiology, 42*, 328–335.

Wittchen, H. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research, 28*(1), 57–84.

Wong, H. M., & Chow, L. Y. (2011). Borderline personality disorder subscale (Chinese version) of the structured clinical interview for DSM-IV axis II personality disorders: A validation study in Cantonese-speaking Hong Kong Chinese. *East Asian Archives of Psychiatry, 21*(2), 52–57.

Wong, S. S., & Fong, K. N. (2009). Reliability and validity of the telephone version of the Cantonese Mini-mental State Examination (T-CMMSE) when used with elderly patients with and without dementia in Hong Kong. *International Psychogeriatrics, 21*(2), 345–353.

Wongchaisuwan, T., Sithinamsuwan, P., Udommongkol, C., & Wongmek, W. (2005). Factors influencing MMSE-T score among Thai subjects. *Journal of the Medical Association of Thailand, 88*(S3), S155–S158.

Wongpakaran, T., Wongpakaran, N., Bookkamana, P., Boonyanaruthee, V., Pinyopornpanish, M., Likhitsathian, S., … Srisutadsanavong, U. (2012). Interrater reliability of Thai version of the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (T-SCID II). *Journal of the Medical Association of Thailand, 95*(2), 264–269.

World Health Organization (WHO). (n.d.). About the WHO CIDI. *The World Mental Health Composite International Diagnostic Interview*. Retrieved from http://www.hcp.med.harvard.edu/wmhcidi/about.php#jump2

Zuckerman, E. L. (2005). *Clinician's thesaurus: The guide to conducting interviews and writing psychological reports*. New York, NY: The Guilford Press.