

A symptom profile of depression among Asian Americans: is there evidence for differential item functioning of depressive symptoms?

Z. Kalibatseva^{1*}, F. T. L. Leong¹ and E. H. Ham²

¹Department of Psychology, Michigan State University, East Lansing, MI, USA

²Department of Counseling, Educational Psychology and Special Education, Michigan State University, East Lansing, MI, USA

Background. Theoretical and clinical publications suggest the existence of cultural differences in the expression and experience of depression. Measurement non-equivalence remains a potential methodological explanation for the lower prevalence of depression among Asian Americans compared to European Americans.

Method. This study compared DSM-IV depressive symptoms among Asian Americans and European Americans using secondary data analysis of the Collaborative Psychiatric Epidemiology Surveys (CPES). The Composite International Diagnostic Interview (CIDI) was used for the assessment of depressive symptoms. Of the entire sample, 310 Asian Americans and 1974 European Americans reported depressive symptoms and were included in the analyses. Measurement variance was examined with an item response theory differential item functioning (IRT DIF) analysis.

Results. χ^2 analyses indicated that, compared to Asian Americans, European American participants more frequently endorsed affective symptoms such as ‘feeling depressed’, ‘feeling discouraged’ and ‘cried more often’. The IRT analysis detected DIF for four out of the 15 depression symptom items. At equal levels of depression, Asian Americans endorsed feeling worthless and appetite changes more easily than European Americans, and European Americans endorsed feeling nervous and crying more often than Asian Americans.

Conclusions. Asian Americans did not seem to over-report somatic symptoms; however, European Americans seemed to report more affective symptoms than Asian Americans. The results suggest that there was measurement variance in a few of the depression items.

Received 3 February 2013; Revised 3 January 2014; Accepted 10 January 2014; First published online 20 February 2014

Key words: Asian American, assessment, culture, depression, differential item functioning.

Introduction

Major depression is a mental disorder that has been identified in all countries and among all ethnic and racial groups studied (Weissman *et al.* 1996; Kessler *et al.* 2003). Major depression is a major public health problem and is projected to be the second largest contributor to global disease burden by 2020 (Murray & Lopez, 1996). The ubiquity and the serious consequences of major depression call for prompt actions in increasing our understanding of its phenomenology, assessment, diagnosis and treatment. Although depression is found universally, it is also thought that culture plays an important role in the aforementioned processes and affects the way depressive symptoms are experienced, expressed, described and measured.

A wide array of theoretical and clinical publications attests to the existence of cultural differences in the expression and experience of depression (e.g. Tanaka-Matsumi & Marsella, 1976; Kleinman & Good, 1985; Marsella, 1987; Kirmayer & Jarvis, 2006). One of the most widely circulated claims in cross-cultural psychopathology is that people of Asian descent somatize psychological distress, and depression in particular. It has been hypothesized that Asians report primarily somatic symptoms and endorse affective symptoms (i.e. sad mood, anhedonia) more rarely, which may explain the low rates of depression reported in some Asian countries (Kleinman, 1982).

Epidemiological studies have shown that socially disadvantaged racial and ethnic minority groups in the USA report significantly lower lifetime risk for depression than non-Latino whites despite higher levels of stress (Breslau *et al.* 2006; Williams *et al.* 2007). In a nationally representative sample from the USA, only 9.1% of Asian Americans were diagnosed with depression (Takeuchi *et al.* 2007) compared to 17.9% of

* Address for correspondence: Ms. Z. Kalibatseva, Department of Psychology, Michigan State University, East Lansing, MI 48224, USA. (Email: kalibats@msu.edu)

non-Latino whites (Breslau *et al.* 2006). The discrepancies in prevalence rates in depression across racial and ethnic groups, especially Asian Americans[†], have raised concerns about the cultural equivalence of depression as a construct.

One of the major problems with ethnocultural variations of depressive disorders is evident in the measurement of depressive experiences. The existing assessments of depressive symptoms may have limited cultural validity that may reduce their clinical utility in non-Western populations (Marsella *et al.* 1985; Kalibatseva & Leong, 2011). In particular, measurement variance may result in misclassification, such that minorities are less likely to meet DSM-IV criteria despite similar levels of an underlying disorder. The symptoms of major depression according to DSM-IV may not be equally culturally sensitive to depressive experience in all populations, that is symptoms may be endorsed differently (Marsella *et al.* 1973). Thus, measurement bias or non-equivalence remains a potential methodological explanation for the lower prevalence of depression among certain ethnocultural groups (Rogler *et al.* 2001).

Differential item functioning (DIF) and the assessment of depression

One way to assess measurement variance in depression symptoms among racial and ethnic minorities is to use DIF. Item response theory (IRT) has been used extensively to detect DIF in educational and psychological testing (Hambleton *et al.* 2011), and has gained popularity in clinical assessment and health and mental health settings (Teresi, 2006). In the case of depression assessment, DIF refers to a situation where the probability of endorsing a particular question about depression differs for individuals who have the same underlying level of depression but belong to different ethnic groups (Breslau *et al.* 2008).

Some studies have explored the endorsement of depressive symptoms among Asian Americans with DIF (e.g. Jang *et al.* 2005; Huang *et al.* 2006). However, few studies have used IRT to detect DIF in diagnostic clinical interviews that assess depression among different ethnic groups (Breslau *et al.* 2008; Chang *et al.* 2008; Uebelacker *et al.* 2009). Only one study has included Asian Americans in a DIF analysis of DSM-IV depression symptoms (Uebelacker *et al.* 2009) and another has included South Koreans and Americans of predominantly European origin (Chang *et al.* 2008). Uebelacker *et al.* (2009) found that Asian Americans were more likely to endorse suicidal

ideation given equal levels of depression severity in comparison to European Americans. At the same time, poor concentration was easier to endorse among Asian Americans than European Americans. The study by Chang *et al.* (2008) revealed that Americans favored depressed mood and thoughts of death and Koreans favored low energy and concentration difficulty.

However, Uebelacker *et al.* (2009) did not find any evidence that somatic symptoms would be endorsed more often among all racial/ethnic minorities than European Americans. Furthermore, although low energy was more easily endorsed by Koreans in the study by Chang *et al.* (2008), none of the other somatic symptoms revealed DIF. Therefore, more research is needed to examine the manifestation of depression among Asian Americans and test the assumption that they endorse somatic symptoms of depression more easily.

Most of the existing literature has investigated depression among Asian Americans in out-patient and community samples and relied on self-report questionnaires. The current study builds on the previous research by using a nationally representative sample of Asian Americans and analyzing data from a structured clinical interview. In this study, systematic analyses were performed to examine the frequencies of discrete depressive symptoms to form a symptom profile of depression among Asian Americans. The symptom profile of depressed Asian Americans was then compared to that of depressed European Americans to detect similarities and differences in the manifestation of depression between the two ethnocultural groups. Based on the existing literature, differences in affective and somatic symptoms were expected. Lastly, the study used IRT to look for DIF in the symptoms.

The aim of the current study was to answer the following questions:

- (1) What depressive symptoms have the highest endorsement rate among Asian Americans?

Hypothesis 1: Asian Americans would have the highest endorsement rates for somatic symptoms (e.g. appetite changes, sleep disturbance and loss of energy).

- (2) What symptoms do Asian Americans endorse more often than European Americans and vice versa?

Hypothesis 2: Asian Americans would endorse somatic depressive symptoms more often (e.g. appetite changes, sleep disturbance, psychomotor disturbance and loss of energy) than European Americans, who would endorse affective depressive symptoms more often (e.g. depressed

[†] The notes appear after the main text.

mood, discouragement, anhedonia and self-reproach) than Asian Americans.

- (3) Is there measurement variance in depression symptoms among Asian Americans and European Americans?

Hypothesis 3: Given equal levels of depression, Asian Americans would endorse somatic symptoms more easily than European Americans and European Americans would endorse affective symptoms more easily than Asian Americans.

The ultimate aim of this study was twofold: to provide information about the nature of depression among Asian Americans and to look for potential measurement variance. A more nuanced understanding of the phenomenology of depression among Asian Americans will allow for better recognition, assessment and diagnosis of depression in this population.

Method

Secondary data analysis was conducted using data from the Collaborative Psychiatric Epidemiology Surveys (CPES; Pennell *et al.* 2004). The CPES includes the National Comorbidity Survey-Replication (NCS-R), the National Latino and Asian American Study (NLAAS) and the National Survey of American Life (NSAL; Heeringa *et al.* 2004). The NLAAS is the first nationally representative community household epidemiological survey of Latinos and Asian Americans in the USA (Alegria *et al.* 2004a). The NCS-R and NSAL are nationally representative surveys of English-speaking household residents aged ≥ 18 years who live in the coterminous United States. These surveys collected data on the lifetime and 12-month prevalence of psychiatric disorders and mental health use. To examine the proposed research questions, data for Asian Americans and European Americans were drawn from CPES. The CPES studies are compatible in the sampling methods and measures they used and allow comparisons of psychiatric disorder characteristics between European, Hispanic, African American and Asian American adults. The CPES dataset is the largest available epidemiological dataset that includes a nationally representative sample of Asian Americans and European Americans.

Sampling design

The CPES used a four-stage stratified probability sampling procedure to recruit and survey adult non-institutionalized Asian Americans and European Americans. The sampling procedures are described in detail in Heeringa *et al.* (2004).

Participants

The participants were 2095 Asian Americans drawn from the NLAAS, 189 Asian Americans and 6696 non-Latino Whites from the NCS-R, and 891 non-Latino Whites from the NSAL. Among the NLAAS respondents there were 600 Chinese, 508 Filipino, 520 Vietnamese, and 467 Other Asians. The Asian American participants from the NCS-R ($n=189$) were part of the Other Asian category.

Measures

The NLAAS measures were translated into four Asian languages: Cantonese, Mandarin, Tagalog and Vietnamese. The core CPES questionnaire was based on the World Health Organization (WHO) version of the Composite International Diagnostic Interview (CIDI 3.0). The CIDI is described in detail in Kessler & Üstün (2004). The WMH-CIDI starts with a screening section that includes questions about particular disorders. If participants endorsed one of the questions for a particular disorder, they were asked to complete the module associated with the disorder later.

This study aimed to explore depressive symptoms among all participants who completed the WMH-CIDI depression module. The rationale for examining this sample was that it would provide more variation within the sample and allow the examination of participants with subthreshold levels of depression who may otherwise be ignored.

Race/ethnicity. The NLAAS targeted Asian American participants of Vietnamese, Filipino and Chinese background. There was a fourth category of 'Other Asian' that included participants of all other Asian ethnicities in the NLAAS and all Asian participants from the NCS-R. The four categories were combined to create one category of Asian Americans. European Americans were drawn from the NCS-R and the NSAL.

Depressive symptoms. The depression section of the WMH-CIDI included questions about depressive symptoms during the most severe major depressive episode (MDE) in the participant's life (V00863–V00892). Questions were relevant to DSM-IV MDE symptoms, and symptoms were rated as present or absent. Based on the DSM-IV MDE diagnostic rules, the endorsement or denial of a symptom (e.g. worthlessness) defined further inquiry about other related symptoms (e.g. guilt).

Lifetime DSM-IV MDE. Lifetime DSM-IV MDE diagnosis was coded as present or absent based on the number of symptoms and rule-outs.

Procedure

Data collection for the NCS-R, NSAL and NLAAS took place between 2001 and 2003. Initially, households and respondents were selected based on probability sampling. The interviewers obtained informed consent and conducted interviews by telephone or in person using computer-based software. All instruments were translated and back-translated into Cantonese, Mandarin, Tagalog and Vietnamese according to standard techniques (Alegria et al. 2004b). Participants received monetary compensation for their participation. Pennell et al. (2004) detail the development and implementation of the CPES studies.

Data analyses

Only participants who endorsed depressive experiences in the screening questions and answered questions about symptoms of their most severe MDE were included in the analyses. Therefore, the subsequent analyses are with a limited sample. Frequencies were obtained and χ^2 tests conducted with the Complex Samples module in IBM SPSS 20.0 (SPSS Inc., USA) using sample weights and controlling for sample design effects due to sample stratification and clustering. Additionally, Mplus 6.0 (<https://www.statmodel.com/>) was used for confirmatory factor analysis; and IRTLRDIF 2.0 (Thissen, 2001) was used for DIF analyses.

Confirmatory factor analysis (CFA). The 29 items used for the previous analyses were reduced to 15 items by combining some of the items based on the CIDI flow. Thus, appetite and weight items were recoded into appetite disturbance; insomnia and hypersomnia into sleep disturbance; and agitation and retardation into psychomotor disturbance. Cognitive difficulty items were combined in one item along with self-reproach items; and five suicidality items were clustered in one suicidality item. Cronbach's α , a measure of the internal consistency of the 15 items, was 0.70. The hypothesized one-common-factor structure for depression was evaluated with a CFA using the Asian and European samples together. As all indicator variables were dichotomous, a weighted least squares means and variance adjusted (WLSMV) approach was applied as the estimator of the polychoric correlations. Studies have shown that WLSMV performs well (Muthén et al. 1997; Flora & Curran, 2004). As a result, the overall model fit was acceptable [$\chi^2=803.15$, $df=90$, root mean square error of approximation (RMSEA)=0.059, Comparative Fit Index (CFI)=0.855], and all factor loadings were above 0.3, indicating that the one common latent factor model was tenable in both samples.

DIF. IRT likelihood ratio (IRT-LR) tests were applied for detecting DIF (Thissen, 2001; Cohen et al. 1996; Ankenmann et al. 1999). The IRT-based procedure for DIF detection is preferred over the observed score-based procedure when the latent trait's distribution is likely to differ between the comparison groups. The basic idea of the IRT-LR is to compare the goodness of fit between two models: an unconstrained and a constrained model. The unconstrained model allows item k 's parameters to be estimated differently between the comparison groups, and the constrained model hypothesizes item k 's parameters to be equal between the groups. When the unconstrained model and the constrained model for item k are statistically different in model fit, item k is considered to have DIF between Asian and European Americans.

The test was carried out in two steps. First, a fully unconstrained model hypothesizing that both slope and threshold parameters of item k were different between the two groups was compared to a fully constrained model hypothesizing that both slope and threshold parameters of item k were identical between the groups. When no difference in model fit between the two models was found, we concluded that the item did not show DIF. If the difference in model fit was statistically significant, we took the second step: a compact unconstrained model with different threshold parameters between the two groups was compared with a compact constrained model with equal threshold parameters under the equal slope parameter. When a difference in model fit between the two compact constrained models was found, we concluded that the item showed DIF. All test statistics and item parameters were obtained from IRTLRDIF 2.0 (Thissen, 2001). More information about this method is provided in Cohen et al. (1996).

Results

Screening

Twenty-three per cent ($n=2284$; 310 Asian Americans and 1974 European Americans) of all Asian and European Americans in the CPES ($n=9871$) endorsed depressed mood, lack of interest or discouragement for a period of at least 2 weeks during their lifetime and completed questions about their depressive symptoms. Thus, 13.2% of Asian Americans compared to 25.9% of European Americans completed questions about their depressive symptoms [$\chi^2(1, n=103)=45.467$, $p<0.001$]. The proportion of Asian Americans with lifetime MDE (9.1%, $s.e.=0.8\%$) was significantly lower than that of European Americans (20.3%, $s.e.=0.5\%$) [$\chi^2(1, n=88)=38.660$, $p<0.001$].

Table 1. Descriptives

Variables	Asian (n=310)	European (n=1974)
Ethnicity (%)		
Vietnamese	9.6 (1.8)	N.A.
Filipino	16.3 (1.5)	N.A.
Chinese	31.1 (3.3)	N.A.
Other Asian	42.9 (2.9)	N.A.
Gender (%)		
Male	38.7 (2.7)	39.2 (1.4)
Female	61.3 (2.7)	60.8 (1.4)
Age (years)	39.22 (0.88)	44.15 (0.65)
Education (%)		
0–11 years	9.7 (2.2)	12.3 (1.0)
12 years	15.3 (1.9)	30.1 (1.2)
13–15 years	25.8 (2.5)	30.2 (0.9)
≥ 16 years	49.2 (2.4)	27.5 (1.3)
Marital status (%)		
Married/cohabiting	53.7 (2.5)	54.6 (1.6)
Divorced/separated	14.2 (1.7)	24.9 (1.1)
Never married	32.2 (1.9)	20.6 (1.6)
MDE diagnosis		
Endorsed (%)	69.0 (2.2)	78.5 (0.9)

MDE, Major depressive episode; N.A., not applicable.
Values given as mean (standard error).

Descriptives

Descriptive statistics are presented in Table 1. The mean age was 39.22 (s.e.=0.88) years for Asian Americans and 44.15 (s.e.=0.65) years for European Americans. Both racial groups had more females than males. Three-quarters (77%) of the participants at risk for depression met criteria for DSM-IV MDE. Of those participants, European Americans (78.3%) were more likely to be diagnosed with MDE than Asian Americans (69%) [$\chi^2(1, n=71)=3.037, p=0.001$].

Question endorsement, symptom prevalence and symptom profile

The endorsement rates of depressive symptoms according to DSM-IV MDE criteria for each racial group and χ^2 comparisons are presented in Table 2. Among the most frequently endorsed symptoms (>70%) for Asian Americans were feeling depressed, feeling discouraged, trouble sleeping, low energy, trouble concentrating, loss of self-confidence and feeling less talkative. Thus, there was partial support for hypothesis 1.

χ^2 analyses: ethnicity \times depressive symptoms

Significant differences in endorsement rates were found between Asian Americans and European

Americans for 10 of the 29 questions, with Asian Americans endorsing all nine questions less frequently than European Americans (see Fig. 1). In particular, Europeans were significantly more likely than Asians to report feeling sad/empty/depressed, feeling discouraged about life, loss of interest, larger appetite, weight gain, loss of self-confidence, feeling worthless, feeling guilty, desire to be alone rather than with friends, and crying often (see Table 2).

IRT-LR tests for detecting DIF (IRT-LR DIF)

As the overall model fit was significantly better in fitting the two-parameter logistic (2PL) model [$-2 \log$ likelihood ($-2LL$)=41056.95] than in fitting the 1PL model ($-2LL$ =41745.09), and the estimated slope parameters were substantially different across items (see Table 3), the 2PL model was applied for the following IRT-LR tests. In addition, as no interaction between group and threshold was hypothesized, this study focused on detecting uniform DIF. Accordingly, the DIF tests were performed only with threshold parameters.

The results are displayed in Table 3. Notably, the estimated means of the latent trait, depression, in the Asian group were consistently lower than those in the European group. The estimated means for Asian Americans ranged from -0.21 to -0.17 and were fixed to 0 for European Americans. Under these distributions, we found that four items (i.e. worthless, nervous, cried often, and appetite change) demonstrated DIF. Specifically, when their estimated depression levels were equal, Asian Americans reported more easily that they felt worthless and experienced appetite changes compared with European Americans. By contrast, European Americans reported feeling nervous and crying more often than Asian Americans given equal levels of depression (see Fig. 2).

Discussion

This study examined a symptom profile of depression among a nationally representative sample of Asian Americans and compared it to that of European Americans. Overall, Asian Americans at risk for depression had high endorsement rates (>70%) for a variety of depressive symptoms, including depressed mood, discouragement, insomnia, loss of energy, trouble concentrating, loss of self-confidence and decreased talkativeness. This pattern is partially consistent with our hypothesis that somatic symptoms would be most prevalent among depressed Asian Americans. Indeed, trouble sleeping and low energy are somatic in nature and the endorsement rate of appetite/weight changes was above 70% when endorsement rates of

Table 2. Endorsement and prevalence of depressive symptoms among Asian Americans (n=310) and European Americans (n=1974)

Symptom item	Asian Americans, mean % (S.E.)	European Americans, mean % (S.E.)	Total	χ^2	p
Depressed mood					
Felt sad, empty or depressed	89.0 (1.7)	93.1 (0.7)	92.9	1.693	0.015
Discouraged about things in your life	80.1 (2.4)	87.2 (0.8)	86.9	2.902	0.003
Anhedonia					
Lose interest in almost all things	68.3 (2.4)	74.8 (1.1)	74.6	1.490	0.014
Appetite/weight change					
Smaller appetite	64.9 (3.6)	64.1 (1.4)	64.2	0.014	0.854
Larger appetite	8.3 (2.1)	14.1 (0.8)	13.9	1.776	0.039
Gain weight	9.7 (2.0)	15.0 (0.8)	14.8	1.400	0.029
Lose weight	55.2 (3.7)	56.8 (1.5)	56.8	0.067	0.676
Sleep disturbance					
Insomnia	74.1 (2.3)	73.3 (1.3)	73.3	0.021	0.763
Sleep a lot more than usual	13.7 (2.0)	16.3 (0.9)	16.2	0.322	0.283
Loss of energy					
Low energy	83.3 (2.7)	83.7 (1.0)	83.7	0.005	0.914
Psychomotor retardation/agitation					
Talk/move more slowly	52.5 (3.1)	51.7 (1.2)	51.8	0.014	0.824
Restless and jittery	10.7 (2.3)	13.1 (0.9)	13.0	0.296	0.384
Cognitive difficulties					
Trouble concentrating	72.9 (3.8)	77.3 (0.8)	77.1	0.676	0.231
Slow thought	55.3 (3.5)	55.2 (1.2)	55.2	0.000	0.973
Indecisive	56.6 (3.3)	62.4 (1.2)	62.3	0.922	0.098
Self-reproach					
Lost self-confidence	71.5 (2.7)	78.3 (1.0)	78.1	1.654	0.016
Feel not as good as other people	61.5 (3.5)	59.3 (1.3)	59.4	0.134	0.542
Felt worthless	59.4 (5.1)	70.2 (1.2)	69.9	2.164	0.031
Feelings of extreme guilt	42.5 (3.3)	51.5 (1.3)	51.2	2.027	0.010
Suicidality					
Thought of death	51.0 (4.0)	58.5 (1.3)	58.3	1.448	0.083
Would be better off dead	43.1 (3.3)	43.3 (1.4)	43.3	0.001	0.958
Thought of committing suicide	26.4 (3.2)	33.4 (1.1)	33.2	1.401	0.058
Made a suicide plan	11.8 (2.1)	11.2 (0.9)	11.2	0.022	0.797
Made a suicide attempt	8.9 (1.5)	8.6 (0.6)	8.6	0.005	0.868
Other psychological problems					
Irritable, grouchy, or in a bad mood	59.9 (3.4)	58.3 (1.0)	58.3	0.063	0.655
Nervous or anxious	59.3 (3.2)	64.6 (1.4)	64.4	0.762	0.134
Alone rather than with friends	67.0 (2.4)	77.0 (0.9)	76.7	3.794	0.000
Less talkative	80.8 (3.0)	81.2 (0.8)	81.2	0.010	0.880
Often in tears	60.4 (2.6)	66.9 (0.9)	66.7	1.113	0.023

S.E., Standard error.

Percentages are weighted.

Statistical tests are design adjusted. Boldface type indicates a significant χ^2 test for difference at the $p=0.05$ level.

decreased (64.9%) and increased (8.3%) appetite were considered jointly.

However, high endorsement rates of affective symptoms were also observed among Asian Americans. In particular, feeling sad, feeling discouraged about things in life and losing self-confidence were endorsed

most often. The high endorsement rates of affective symptoms could have several possible explanations. First, the results are consistent with findings that Asian Americans experience depressive affective symptoms and are not denying the underlying psychological problems in depression (e.g. Lu *et al.* 2010).

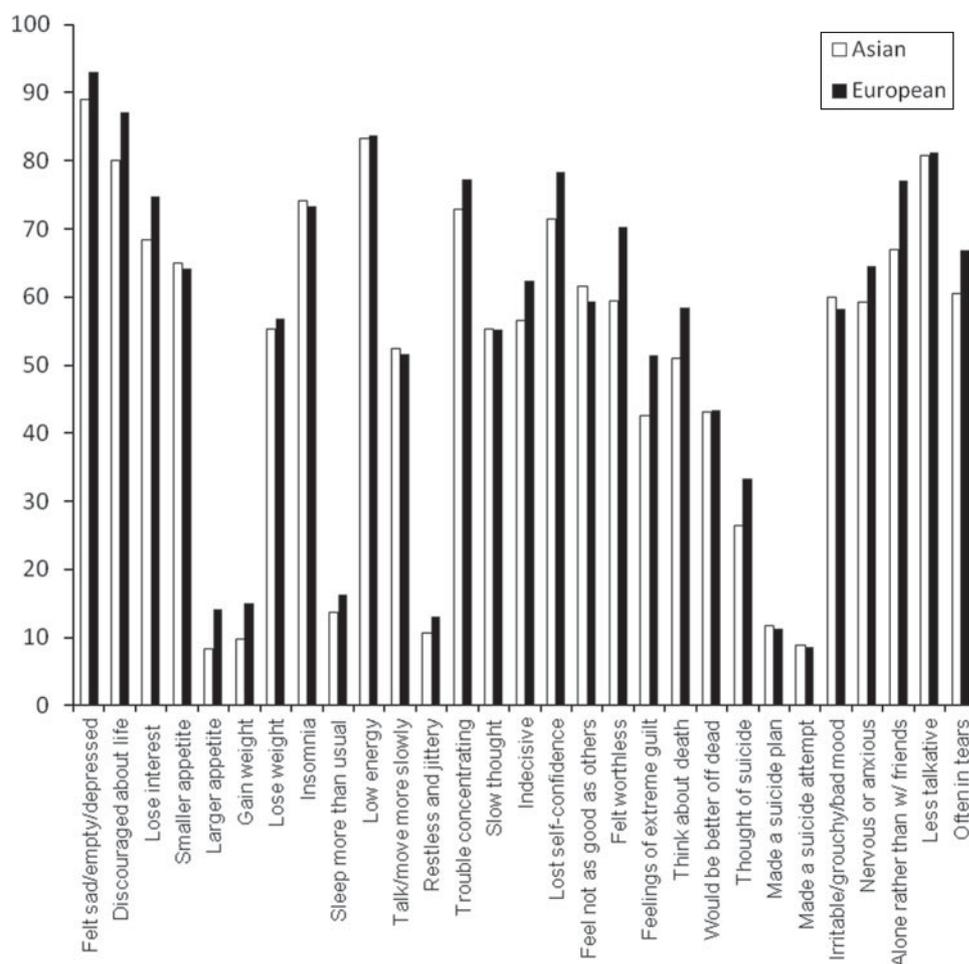


Fig. 1. Frequency of depressive symptoms among Asian Americans and European Americans.

Second, the high endorsement rates of affective symptoms among Asian Americans might be related to their acculturation level because more acculturated Asian Americans may be more likely to manifest depression similarly to European Americans (Lee, 2002). Third, the diagnostic measure (WMH-CIDI) in the CPES used a screening process that emphasized affective symptoms of depression (i.e. felt sad, empty, or depressed; felt discouraged; lost interest in things). Therefore, it is possible that the participants who completed the WMH-CIDI depression module were already primed to report their affective symptoms.

The screening process that was used in CPES limited our sample to 23% of the original sample who answered all questions about their most severe depressive episode. Although these results suggest that the core features of depression are present in different racial/ethnic groups, it is important to consider this finding in the context of the methodology that was used, as noted in the following classic example from cross-cultural psychology. The International Pilot Study of Schizophrenia (IPSS) conducted by the WHO in the

1970s provided the first data for cross-cultural comparisons of schizophrenia and concluded that psychoses presented similarly across cultures. However, Kleinman (1988) later criticized the methodology of the IPSS for using stringent inclusion and exclusion criteria, which yielded an 'artificially homogenous sample' (Thakker & Ward, 1998, p. 516), and suggested that the observed 'similarity was an artifact of methodology' (Kleinman, 1988, p. 19). In the present study, a similar phenomenon may have been observed, as the screening questions for depression were based on DSM-IV MDE criteria. Specifically, the CIDI may show good diagnostic validity in undeprived contexts within the USA but may miss cases of depression in deprived environments (Hicks, 2002). Nevertheless, considering the scarcity of research with racial and ethnic minorities, we consider that this study provides important information about the expression and experience of depression.

To examine similarities and differences in depressive symptoms among Asian Americans and European Americans, the frequencies of endorsement were

Table 3. Estimated item parameters by group and IRT-LT DIF test results

	DIF test ^a			European (E) (reference group)		Asian (A) (focal group)	
	Model comparison	G2 (df)	Group comparison	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
Depressed	Full	2.4 (2)		1.74	-2.16	1.34	-2.33
Discouraged	Full	4* (2)		1.55	-1.69	1.31	-1.6
	Compact	3.2 (1)		1.50	-1.72	1.50	-1.49
Lost interest	Full	0.2 (2)		1.51	-1	1.55	-0.94
Low energy	Full	2.5 (2)		1.44	-1.4	1.85	-1.35
Concentration	Full	2.4 (2)		1.62	-1.31	2.02	-1.31
Worthless	Full	4.6* (2)		1.15	-0.31	1.32	-0.53
	Compact	4.0* (1)	A > E	1.17	-0.30	1.17	-0.56
Death/suicide	Full	2.3 (2)		0.89	1.02	0.77	1.42
Irritable	Full	3.3 (2)		0.93	-0.36	1.09	-0.55
Nervous	Full	4.5* (2)		0.93	-0.7	1.03	-0.38
	Compact	4.2* (1)	E > A	0.95	-0.69	0.95	-0.39
Alone	Full	4.6* (2)		1.23	-1.15	1.00	-1.03
	Compact	3.3 (1)		1.19	-1.18	1.19	-0.94
Less talkative	Full	0.6 (2)		1.39	-1.31	1.45	-1.37
Cried often	Full	5.3* (2)		0.67	-1.16	0.54	-0.85
	Compact	4.5* (1)	E > A	0.65	-1.19	0.65	-0.75
Appetite change	Full	4.6* (2)		0.9	-1.41	1.01	-1.64
	Compact	4.2* (1)	A > E	0.92	-1.39	0.92	-1.75
Sleep disturbance	Full	3 (2)		1.56	-1.65	1.18	-2.05
Psychomotor	Full	2.9 (2)		1.28	-0.48	1.76	-0.45

IRT-LT DIF, Item response theory likelihood ratio differential item functioning.

Boldface type indicates items that revealed differential item functioning.

^aIn the model comparison, 'full' means that a constrained model with equal *a* and *b* parameters between the two groups was compared to an unconstrained model with group-specific *a* and *b* parameters; 'compact' means a constrained model with an equal *b* parameter was compared with an unconstrained model with a different *b* parameter between the two models: $G2(df) = -2(LL_{\text{constrained}} - LL_{\text{unconstrained}})$.

compared. Ten differences were found and, for all of them, European Americans endorsed the symptoms significantly more often than Asian Americans. In particular, European Americans more frequently endorsed feeling sad, empty or depressed, feeling discouraged about things in life, losing interest in almost all things, having a larger appetite, gaining weight, losing self-confidence, feeling worthless, feeling extreme guilt, wanting to be alone, and crying often. These findings were only partially consistent with our second hypothesis, which predicted that Asian Americans would endorse somatic symptoms more often and European Americans would be more likely to endorse affective symptoms. We observed that, although the first part of this hypothesis was not supported, the second part was confirmed. European Americans were more likely to endorse depressed mood, anhedonia and self-reproach symptoms. This finding is consistent with the proposition that Westerners are more likely to 'psychologize' depression than Asians whereas somatic

complaints are ubiquitous (Kirmayer & Young, 1998; Ryder *et al.* 2008).

Additionally, higher endorsement rates of increased appetite and weight gain were observed among European Americans but not among Asian Americans. Although we did not have a hypothesis for this particular symptom, Marsella *et al.* (1973) reported a similar pattern, with Japanese and Chinese Americans endorsing a lack of appetite and European Americans endorsing an urge to eat.

The DIF analysis indicated that four of the 15 items examined revealed measurement variance. Assuming equal levels of depression, Asian Americans endorsed feeling worthless and appetite disturbances more easily whereas European Americans endorsed feeling nervous and crying often more easily. These results provide partial support for the third hypothesis, as Asian Americans endorsed one somatic and one affective symptom more easily than European Americans. Consistent with the hypothesis and

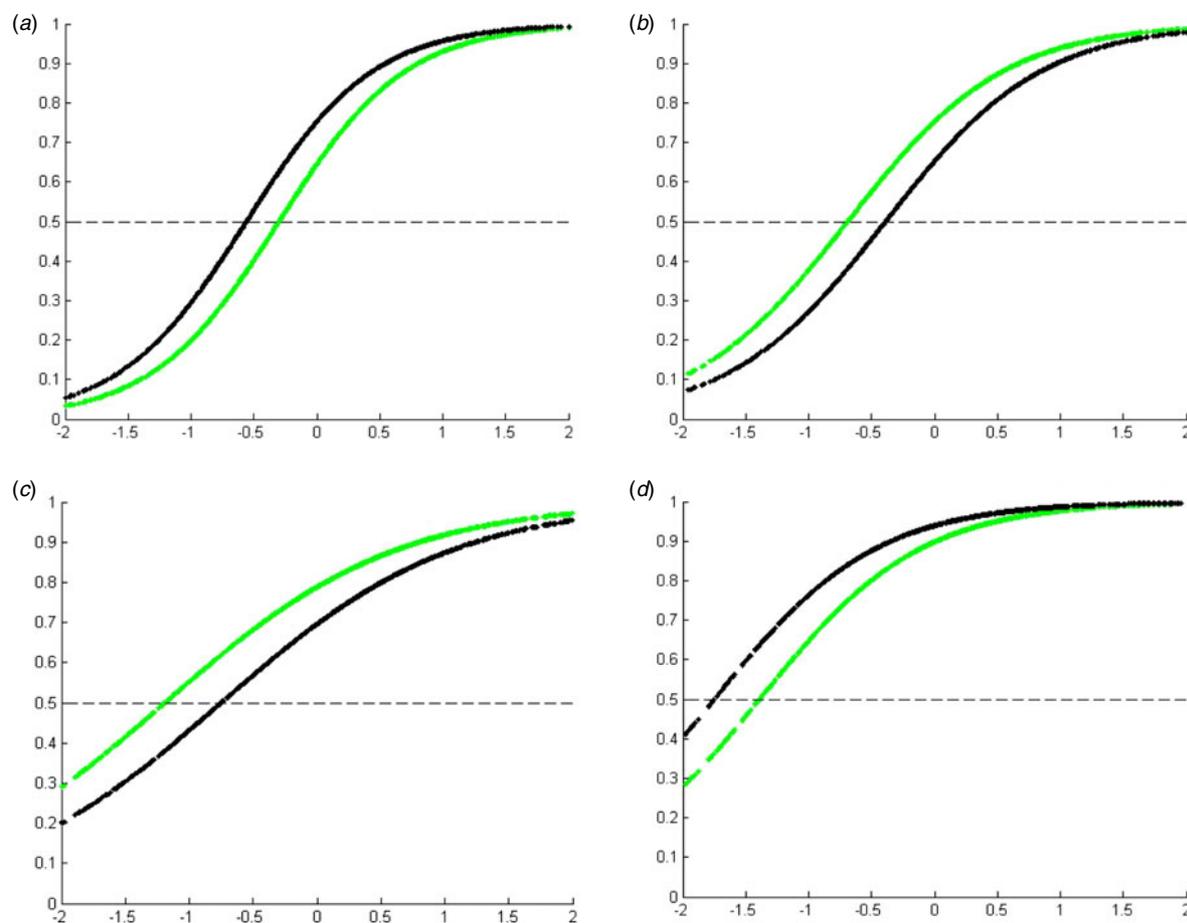


Fig. 2. Expected probability of endorsing through the estimated level of depression by group membership (dark lines represent Asian Americans; light lines represent European Americans; the t statistic is associated with the difference in the expected probability between the two groups). (a) Worthless ($t=1.47$, $p=0.142$); (b) Nervous ($t=-11.44$, $p<0.001$); (c) Cried often ($t=-13.38$, $p<0.001$) (d) Appetite ($t=0.42$, $p=0.671$).

previous research, European Americans reported two affective symptoms more easily than Asian Americans with the same level of depression (Ryder *et al.* 2008).

We can speculate that the expression of negative emotions is more socially acceptable among European Americans than Asian Americans due to possible differences in the interaction between self-construal and emotional regulation. Independent self-construal involves construing the self as an individual, whose behavior is organized and meaningful based on the person's own feelings, thoughts and actions. Interdependent self-construal entails perceiving oneself as part of social relationships and realizing that one's behavior is contingent on and organized by the person's perception of others' thoughts, feelings and behaviors (Markus & Kitayama, 1991). Markus & Kitayama (1991) suggested that independent self-construal is observed in Western cultures and interdependent self-construal is more common in Asian cultures, although variations within cultures are possible. In terms of emotional regulation, the authors implied that a

person's self-construal can affect the expression, intensity and frequency of emotions. Specifically, those with independent selves learn how to communicate their 'ego-focused' emotions, such as sadness or frustration, very effectively. By contrast, people with interdependent self-construal need to control and de-emphasize their private feelings so as to fit into the interpersonal context. Therefore, European Americans may put more weight on expressing negative affect (e.g. depressed mood, discouragement, crying often) than Asian Americans when they suffer from depression. Using the self-construal framework, we can also discuss European Americans' higher endorsement of wanting to be alone rather than with friends. The desire to be on one's own when depressed may be related to the concept of independence and individualism seen more often among Western cultures. Conversely, Asian Americans may seek help from their social network or they may simply not have the choice to be on their own because isolation and avoidance are not socially appropriate.

The DIF results suggest evidence for measurement non-equivalence for two affective items for European Americans and for one somatic and one affective item for Asian Americans. However, the items that demonstrated measurement variance would probably not explain the significantly higher rate of depression diagnosis among European Americans than Asian Americans.

Clinical implications

The reported findings have important implications for mental health professionals and primary care physicians who work with Asian Americans. A culturally informed assessment of depressive symptoms among Asian Americans would equally emphasize affective and somatic symptoms and may inquire about self-worth and appetite changes. Although affective symptoms may be present among depressed Asian Americans, they may not be the most salient ones or the reason why clients sought help in the first place. In addition, taking into consideration the client's acculturation and probing further may be essential for the assessment, diagnosis and treatment of depression among Asian Americans.

Limitations

This study had several limitations that need to be addressed in future research. First, the screening process that was used in the WMH-CIDI limited our sample to Asian Americans who reported past experiences of sad, empty or depressed mood, discouragement, and lack of interest. Although this screening process is consistent with the DSM-IV diagnosis, it might have eliminated Asian American participants who experienced depression differently. However, we consider that the sample we examined still presented with reasonable cultural variations in depression and provided valid data to answer our research questions. Another limitation of the current study was the small cell counts to examine differences in depressive symptoms based on ethnicity. Although one of the strengths of the CPES dataset is the oversampling of Chinese, Filipino and Vietnamese participants, we could not take advantage of this feature because of the small number of individuals in each ethnic group.

Finally, considerable research has shown that acculturation level is a significant moderator of Asian Americans' mental health status and service utilization (Leong & Lau, 2001). It is possible that measurement variance in depression may differentially affect Asian Americans with high *versus* low levels of acculturation. However, acculturation was not measured in this dataset.

Directions for future research

The systematic examination of symptom profiles of mental disorders among ethnic and racial minorities and cross-culturally can provide valuable information for improving assessment, diagnosis and treatment. In the case of depression, it may be particularly important to examine symptoms among individual ethnic groups of Asian Americans and other racial minorities. In addition, although we found cross-racial differences, testing the mechanisms behind these differences in depressive symptoms among Asian Americans will be an important next step. As mentioned earlier, another topic that deserves more attention is the effect of acculturation on the overall prevalence of depression, and of depressive symptoms in particular.

Furthermore, given the importance of face and shame in Asian cultures (Zane & Yeh, 2002), differential symptom expression of depression in questionnaires (e.g. the Beck Depression Inventory-II) *versus* clinical interviews (e.g. the CIDI) could be examined in future studies to determine possible method variance (Ryder *et al.* 2008). Lastly, future research of depressive symptoms endorsed by both depressed and non-depressed Asian Americans may help to elucidate the diagnostic validity of the DSM-5 MDE diagnosis. In general, having a sample with a wide variation of clinical symptoms, as opposed to a sample limited to participants with a DSM diagnosis, will allow us to examine cultural variations of depression and other disorders more thoroughly.

Acknowledgments

We thank Dr N. Schmitt for his help with earlier versions of this manuscript. We also thank the CPES principal investigators, M. Alegria, J. Jackson, R. Kessler and D. Takeuchi, for initiating and completing this important project and allowing other researchers to use the datasets.

Declaration of Interest

None.

Notes

- ¹ When referring to Asian Americans, we include immigrants from Asian countries to the USA (first generation) and Americans of Asian descent (second, third or fourth generation).

References

Alegria M, Takeuchi DT, Canino G, Duan N, Shrout P, Meng XL, Vega W, Zane N, Vila D, Woo M, Vera M,

- Guarnaccia P, Aguilar-Gaxiola S, Sue S, Escobar J, Lin KM, Gong F** (2004a). Considering context, place, and culture: the National Latino and Asian American Study. *International Journal of Methods in Psychiatric Research* **13**, 208–220.
- Alegria M, Vila D, Woo M, Canino G, Takeuchi D, Vera M, Febo V, Guarnaccia P, Aguilar-Gaxiola S, Shrout P** (2004b). Cultural relevance and equivalence in the NLAAS instrument: integrating etic and emic in the development of cross-cultural measures for a psychiatric epidemiology and services study of Latinos. *International Journal of Methods in Psychiatric Research* **13**, 270–288.
- Ankenmann RD, Witt EA, Dunbar SB** (1999). An investigation of the power of the likelihood ratio goodness-of-fit statistic in detecting differential item functioning. *Journal of Educational Measurement* **36**, 277–300.
- Breslau J, Aguilar-Gaxiola S, Kendler KS, Su M, Williams D, Kessler RC** (2006). Specifying race-ethnic differences in risk for psychiatric disorder in a USA national sample. *Psychological Medicine* **36**, 57–68.
- Breslau J, Javaras KN, Blacker D, Murphy JM, Normand ST** (2008). Differential item functioning between ethnic groups in the epidemiological assessment of depression. *Journal of Nervous and Mental Disease* **196**, 297–306.
- Chang SM, Hahm BJ, Lee JL, Shin MS, Jeon HJ, Hong JP, Lee HB, Lee DW, Cho MJ** (2008). Cross-national difference in the prevalence of depression caused by the diagnostic threshold. *Journal of Affective Disorders* **106**, 159–167.
- Cohen AS, Kim S, Wollack JA** (1996). An investigation of the likelihood ratio test for detection of differential item functioning. *Applied Psychological Measurement* **20**, 15–26.
- Flora DB, Curran PJ** (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* **9**, 466–491.
- Hambleton RK, Bartram D, Oakland T** (2011). Technical advances and guidelines for improving testing practices. In *IAAP Handbook of Applied Psychology* (ed. P. R. Martin, F. M. Cheung, M. C. Knowles, M. Kyrios, L. Littlefield, J. B. Overmier and J. M. Prieto), pp. 338–361. Wiley-Blackwell: Hoboken, NJ.
- Heeringa SG, Wagner J, Torres M, Duan N, Adams T, Berglund P** (2004). Sample design and sampling methods for the Collaborative Psychiatric Epidemiology Studies (CPES). *International Journal of Methods in Psychiatric Research* **13**, 221–240.
- Hicks MHR** (2002). Validity of the CIDI probe flow chart for depression in Chinese American women. *Transcultural Psychiatry* **39**, 434–451.
- Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL** (2006). Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *Journal of General Internal Medicine* **21**, 547–552.
- Jang Y, Kim G, Chiriboga D** (2005). Acculturation and manifestation of depressive symptoms among Korean-American older adults. *Aging and Mental Health* **9**, 500–507.
- Kalibatseva Z, Leong FTL** (2011). Depression among Asian Americans: review and recommendations. *Depression Research and Treatment* **2011**, 320902.
- Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS** (2003). The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association* **289**, 3095–3105.
- Kessler RC, Üstün TB** (2004). The World Mental Health Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research* **13**, 93–121.
- Kirmayer L, Jarvis GE** (2006). Depression across cultures. In *The American Psychiatric Publishing Textbook of Mood Disorders* (ed. D. J. Stein, D. J. Kupfer and A. F. Schatzberg), pp. 699–715. American Psychiatric Publishing: Arlington, VA.
- Kirmayer LJ, Young A** (1998). Culture and somatization: clinical, epidemiological, and ethnographic perspectives. *Psychosomatic Medicine* **60**, 420–430.
- Kleinman A** (1982). Neurasthenia and depression: a study of somatization and culture in China. *Culture, Medicine, and Psychiatry* **6**, 117–190.
- Kleinman A** (1988). *Rethinking Psychiatry: From Cultural Category to Personal Experience*. Free Press: New York, NY.
- Kleinman A, Good B** (1985). *Culture and Depression*. University of California Press: Berkeley, CA.
- Lee S** (2002). Socio-cultural and global health perspectives for the development of future psychiatric diagnostic systems. *Psychopathology* **35**, 152–157.
- Leong FTL, Lau ASL** (2001). Barriers to providing effective mental health services to Asian Americans. *Mental Health Services Research* **3**, 201–214.
- Lu A, Bond MH, Friedman M, Chan C** (2010). Understanding cultural influences on depression by analyzing a measure of its constituent symptoms. *International Journal of Psychological Studies* **2**, 55–70.
- Markus HR, Kitayama S** (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychological Review* **98**, 224–253.
- Marsella AJ** (1987). The measurement of depressive experience and disorder across cultures. In *The Measurement of Depression* (ed. A. J. Marsella, R. M. A. Hirschfeld and M. M. Katz), pp. 376–397. Guilford Press: New York, NY.
- Marsella AJ, Kinzie D, Gordon P** (1973). Ethnic variations in the expression of depression. *Journal of Cross-Cultural Psychology* **4**, 435–458.
- Marsella AJ, Sartorius N, Jablensky A, Fenton F** (1985). Cross-cultural studies of depressive disorders: an overview. In *Culture and Depression* (ed. A. Kleinman and B. Good), pp. 299–324. University of California Press: Berkeley, CA.
- Murray CJ, Lopez AD** (1996). *The Global Burden of Disease. A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020*. Harvard School of Public Health: Cambridge, MA.

- Muthén BO, du Toit SHC, Spisic D** (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes (www.statmodel.com/wlscv.shtml). Accessed 20 January 2014.
- Pennell BE, Bowers A, Carr D, Chardoul S, Cheung GQ, Dinkelmann K, Gebler N, Hansen SE, Pennell S, Torres M** (2004). The development and implementation of the National Comorbidity Survey Replication, the National Survey of American Life, and the National Latino and Asian American Survey. *International Journal of Methods in Psychiatric Research* **13**, 241–269.
- Rogler LH, Mroczek DK, Fellows M, Loftus ST** (2001). The neglect of response bias in mental health research. *Journal of Nervous and Mental Disease* **189**, 182–187.
- Ryder AG, Yang J, Zhu X, Yao S, Yi J, Heine S, Bagby RM** (2008). The cultural shaping of depression: somatic symptoms in China, psychological symptoms in North America? *Journal of Abnormal Psychology* **117**, 300–313.
- Takeuchi DT, Hong S, Gile K, Alegria M** (2007). Developmental contexts and mental disorders among Asian Americans. *Research in Human Development* **4**, 49–69.
- Tanaka-Matsumi J, Marsella AJ** (1976). Cross-cultural variation in the phenomenological experience of depression: I. Word association studies. *Journal of Cross-Cultural Psychology* **7**, 379–396.
- Teresi JA** (2006). Different approaches to differential item functioning in health applications. Advantages, disadvantages and some neglected topics. *Medical Care* **44**, S152–S170.
- Thakker J, Ward T** (1998). Culture and classification: the cross-cultural application of the DSM-IV. *Clinical Psychology Review* **18**, 501–528.
- Thissen D** (2001). IRTLRDIF v. 2.0b: software for the computation of the statistics involved in item response theory likelihood-ratio tests for differential item functioning (www.unc.edu/~dthissen/dl.html). Accessed 17 December 2013.
- Uebelacker LA, Strong D, Weinstock LM, Miller IW** (2009). Use of item response theory to understand differential item functioning of DSM-IV major depression symptoms by race, ethnicity, and gender. *Psychological Medicine* **39**, 591–601.
- Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, Joyce PR, Karam EG, Lee CK, Lellouch J, Lepine JP, Newman SC, Rubio-Stipec M, Wells JE, Wickramaratne PJ, Wittchen H, Yeh EK** (1996). Cross-national epidemiology of major depression and bipolar disorder. *Journal of the American Medical Association* **276**, 293–299.
- Williams DR, Gonzalez HM, Neighbors H, Nesse R, Abelson JM, Sweetman J, Jackson JS** (2007). Prevalence and distribution of major depressive disorder in African Americans, Caribbean blacks and non-Hispanic whites: results from the National Survey of American Life. *Archives of General Psychiatry* **64**, 305–315.
- Zane N, Yeh M** (2002). The use of culturally-based variables in assessment: studies on loss of face. In *Asian American Mental Health: Assessment Theories and Methods* (ed. K. S. Kurasaki, S. Okazaki and S. Sue), pp. 123–137. Kluwer Academic/Plenum Publishers: New York, NY.