Asian American Journal of Psychology
2019, Vol. 10, No. 1, 1–10


1948-1985/19/$12.00 http://dx.doi.org/10.1037/aap0000115

# Measurement Invariance of the Brief Symptom Inventory-18 (BSI-18) Across Asian American Ethnic, Nativity, and Gender Groups

Yun Lu
University of Maryland

Alvin N. Alvarez
San Francisco State University

Matthew J. Miller
University of Maryland

This study was the first to examine the factor structure, measurement invariance, and criterion-related evidence for the construct validity of the Brief Symptom Inventory-18 (BSI-18; Derogatis, 2000) in an Asian American sample. Multigroup confirmative factor analyses were carried out in a community sample of adults of Asian ethnic descent ($N = 624$; $M = 41.47$ years, $SD = 13.28$) across ethnicity (Chinese, Filipinx, and Vietnamese), nativity status (U.S.-born and Asian-born), and gender (female and male) to test a priori competing models and measurement invariance. Results of bifactor analyses and ancillary estimates provided strong evidence for the general factor of the BSI-18. Invariance tests indicated that BSI-18 items were interpreted in a similar fashion, and responses could be meaningfully compared across ethnicity, gender, and nativity samples. In addition, theory-consistent correlations with self-esteem and racial microaggression experiences provided criterion-related evidence for the BSI-18. Results may not generalize to other Asian American groups (e.g., South Asians, Asian elders, and clinical samples), and researchers are strongly encouraged to conduct further internal structure analyses of the BSI-18 among Asian American populations. In conclusion, we recommend calculation and interpretation of the BSI-18 total score as a measure of general psychological distress for Asian American populations.

***What is the public significance of this article?***
This study examined the measurement equivalence and criterion-related evidence for the construct validity of the Brief Symptom Inventory-18 in an Asian American sample. Findings suggest that the total score of the Brief Symptom Inventory-18 is appropriate for use with several Asian American subpopulations across gender and nativity groups as a measure of general psychological distress.

*Keywords:* psychological distress, Asian American, measurement invariance, construct validity

An all-too-common misconception of the Asian American population, referred to as the model minority myth, is that they are well adjusted and thriving in the United States as compared with other racial minority populations (Yoo, Miller, & Yip, 2015). This stereotype obscures the experiences of this heterogeneous population, such as acculturative stress and experiences of racism and discrimination that have a negative impact on their mental health (Alvarez, Juang, & Liang, 2006; Miller, Alvarez, Li, Chen, & Iwamoto, 2016; Miller, Kim, & Benet-Martínez, 2011). In fact, emerging empirical evidence suggests that mental health problems are a bigger concern for Asian Americans than what was presumed (Zane, 2007). For example, Asian Americans reported the highest amount of psychological distress among all racial groups in a national clinical college student sample (Kearney, Draper, & Barón, 2005). There is a need for developing and validating psychometrically sound measures of psychological distress for use with Asian American populations (Wong, Wu, Guo, Lam, & Snowden, 2012). To advance research in this area, we tested the construct validity of the Brief Symptom Inventory-18 (BSI-18; Derogatis, 2000) across ethnicity (Chinese, Filipinx, and Vietnamese), nativity status (US-born and Asian-born), and gender (female and male) groups by examining factor structure, measurement invariance, and criterion-related evidence.

The BSI-18 is a widely used short screening tool for psychological distress and has received much attention in the assessment literature. A principal component analysis with a community sample of 1,134 participants (Derogatis, 2000) produced a three-component model: somatization (SOM), depression (DEP), and

This article was published Online First April 23, 2018.

Yun Lu, Department of Counseling, Higher Education, and Special Education, University of Maryland; Alvin N. Alvarez, College of Health and Social Sciences, San Francisco State University; Matthew J. Miller, Department of Counseling, Higher Education, and Special Education, University of Maryland.

This work was supported in part by the National Institute of Mental Health's Minority Research Infrastructure Program (R24-MH061573-04).

Correspondence concerning this article should be addressed to Yun Lu, Department of Counseling, Higher Education, and Special Education, University of Maryland, 3214 Benjamin Building, 3942 Campus Drive, College Park, MD 20742. E-mail: ylu424@umd.edu

anxiety (ANX). The three-factor model was replicated in an independent study of 8,945 predominantly White cancer survivors (Recklitis et al., 2006). Contrary to these findings, a one-factor model of the BSI-18 was supported in studies of 1,115 Latina American mothers (Prelow, Weaver, Swenson, & Bowman, 2005) and 923 Mexican adults (Torres, Miller, & Moore, 2013). To further complicate matters, a study of 4,711 ethnically diverse women found that the three-factor model was most appropriate for Black and White participants, whereas the one-factor model was determined to be most appropriate for Latina participants (Wiesner et al., 2010). Collectively, these results suggest possible racial/cultural factors that moderate the dimensionality of the BSI-18.

However, although the Brief Symptom Inventory and BSI-18 have been used in Asian American samples (Liu & Suyemoto, 2016), no study has examined the factor structure of the BSI-18 in Asian American populations. The lack of research in this area is problematic. Asian and Asian American clinical samples tend to report higher rates of somatic symptoms and lower rates of affective symptoms compared with European American populations (Ryder et al., 2008; Yen, Robins, & Lin, 2000), but extant BSI-18 psychometric evidence was predominantly based on European American samples. One potential explanation for the unique experience and/or expression of psychological distress in Asian American populations is that Asian cultural beliefs and values, such as holistic views of body, mind, and spirit; forbearance coping; an orientation toward others; and filial piety, may lead to culturally unique symptom expression, including less distinction between physical and mental health symptoms, less endorsement of psychological needs, and more endorsement of physical needs (Iwamasa & Hilliard, 1999; Kalibatseva, Leong, Ham, Lannert, & Chen, 2017).

The Asian American populations comprise many diverse ethnic groups, each of which may have unique endorsement and appraisal of psychological distress due to differences in sociopolitical and immigration histories (e.g., preimmigration trauma, refugee status, and colonialism), socioeconomic status (SES), educational attainment, cultural values, linguistic background, and religion (Kim, Yang, Atkinson, Wolfe, & Hong, 2001; Takeuchi et al., 2007). For example, lower socioeconomic and employment status, compared with other Asian ethnic groups, partly identified Filipinx Americans as a high-risk group for mental disorders (Sanchez & Gaw, 2007). Failure to recognize the uniqueness of specific Asian American ethnic groups contributes to a homogenized view of Asian Americans. Evidence also suggests that nativity and generational statuses, as important predictors of cultural exposure and acculturation, are meaningful correlates of Asian American health and mental health (Takeuchi et al., 2007). For example, Asian-born individuals have been found to have significantly lower rates of mental health disorders than U.S.-born Asian Americans (Hong, Walton, Tamaki, & Sabin, 2014). In terms of gender and psychological distress, U.S.-born Asian American women may be at greater risk for mood disorders, whereas U.S.-born Asian American men may be at greater risk for substance use disorders (Hong et al., 2014). However, much of the group comparisons have been based on analysis of variance (ANOVA) and *t* test rather than latent mean analyses and were susceptible to threats of measurement errors as a result.

## Present Study

The present study tested (a) competing models of the BSI-18 factor structure; (b) the measurement invariance of the BSI-18 across Asian American ethnic, nativity, and gender groups; and (c) criterion-related evidence of BSI-18 in Asian American samples. Specifically, we compared the fit of three a priori competing models: a one-factor model in which all 18 items loaded onto a common psychosomatic distress factor (Prelow et al., 2005; Torres et al., 2013; Wiesner et al., 2010), the aforementioned theory-derived three-factor model in which 18 items loaded onto three correlated factors (Derogatis, 2000; Recklitis et al., 2006; Wiesner et al., 2010), and a bifactor model in which 18 items simultaneously loaded onto three orthogonal group factors and a general factor. After selecting a theoretically consistent and psychometrically sound factor structure, we examined measurement invariance with increasingly stringent tests that determined if the patterns and magnitude of factor loadings, item intercepts, and factor means differed across ethnicity, nativity status, and gender groups (French & Finch, 2006; Vandenberg & Lance, 2000).

For criterion-related evidence, we expected BSI-18 scores to have a negative correlation with self-esteem, given that many theories of depression postulate low self-esteem as a defining feature of depression (Beck, 1967), and the robust relationship between low self-esteem and depression documented in the literature (Joiner, Katz, & Lew, 1999; Kernis, Grannemann, & Mathis, 1991; Orth, Robins, & Roberts, 2008). Because depression was only one factor in the original BSI-18, we expected the relationship to be small to moderate. We also expected a moderate positive relationship between BSI-18 and racial microaggression experiences, because detrimental effects of race-related stress on physical and mental health for people of color, and Asian Americans specifically, have been robustly found in the literature (Clark, Anderson, Clark, & Williams, 1999; Liu & Suyemoto, 2016; Miller, Yang, Farrell, & Lin, 2011; Okazaki, 2009).

## Method

### Participants

Participants of Vietnamese ($n = 209$), Filipinx ($n = 197$), and Chinese ($n = 219$) descent were recruited as part of a larger study (Table 1). All samples had balanced gender representation and a diverse age, education, and SES representation. ANOVAs with Bonferroni post hoc tests and chi-square tests with standardized residual estimates (*SR*) were conducted to examine similarities and differences in demographic representations across ethnic samples. *SR*s greater than $\pm 2$ identify those cells with statistically significant differences between the observed and expected values for a significant chi-square test. Three samples had similar gender representation, $\chi^2(2, 0.5) = 5.29$, $p = .07$. Participants in the Vietnamese sample, on average, had a significantly shorter length of stay in the United States than those in the Chinese and the Filipinx samples by 7–8 years, $p$s $< .001$. In terms of nativity status, Vietnamese participants were overrepresented to be non-U.S.-born, and underrepresented to be U.S.-born, whereas Chinese and Filipinx participants were significantly overrepresented to be U.S.-born, $\chi^2(2, 0.5) = 106.59$, $p < .001$, $|SR\text{s}| > 2.1$. In terms of age,

Table 1
*Sample Descriptive Statistics (N = 625)*

| Characteristics | *n* (%) | | |
|---|---|---|---|
| | Vietnamese (*n* = 209) | Filipinx (*n* = 197) | Chinese (*n* = 219) |
| Age, *M* (*SD*) | 45.8 (13.7) | 39.6 (12.6) | 39.1 (12.5) |
| Nativity | | | |
| Asian-born | 192 (91.9) | 89 (45.2) | 125 (57.1) |
| U.S.-born | 16 (7.7) | 104 (52.8) | 94 (42.9) |
| Missing | 1 (.5) | 4 (2.0) | — |
| Gender | | | |
| Female | 114 (54.5) | 90 (45.7) | 123 (56.2) |
| Male | 94 (45.0) | 107 (54.3) | 96 (43.8) |
| Missing | 1 (.5) | — | — |
| Ethnicity | | | |
| Vietnamese/Filipinx/Chinese | 176 (84.2) | 162 (82.2) | 179 (81.7) |
| Multiracial/ethnic | 14 (6.7) | 32 (16.2) | 12 (5.5) |
| Other | 19 (9.1) | 2 (1.0) | 28 (12.8) |
| Missing | | 1 (.5) | |
| Year of residence, *M* (*SD*) | 22.2 (7.6) | 30.9 (11.0) | 28.9 (12.8) |
| Education | | | |
| Less than high school | 20 (9.7) | — | 14 (6.4) |
| High school | 33 (15.8) | 25 (12.7) | 24 (11.0) |
| College | 121 (57.9) | 127 (64.5) | 113 (51.6) |
| Professional/Graduate | 32 (15.3) | 43 (21.8) | 67 (30.6) |
| Missing | 3 (1.4) | 2 (1.0) | 1 (.5) |
| Socioeconomic status | | | |
| Lower-lower middle | 49 (23.9) | 60 (30.4) | 55 (25.1) |
| Middle | 126 (60.3) | 103 (52.3) | 128 (58.4) |
| Upper middle-upper | 30 (14.4) | 30 (15.2) | 36 (16.5) |
| Missing | 4 (1.9) | 4 (2) | — |

Vietnamese participants were significantly older than Chinese and Filipinx participants by about 7 years, *p*s $< .001$. Although there was an overall significantly different pattern of SES representation across ethnicity, $\chi^2(8, 0.5) = 20.381$, $p = .009$, where Filipinx participants seemed to be overrepresented in lower middle class and underrepresented in upper class, and that Vietnamese participants seemed to be underrepresented in lower middle class and overrepresented in upper class, the specific differences were not statistically significant, as judged by the *SR*s. With regard to educational background, Vietnamese participants were significantly overrepresented with elementary education, $SR = 3.0$, and underrepresented with professional/graduate education, $SR = -2.2$. Chinese participants were significantly overrepresented with professional/graduate education, $SR = 2.4$.

We compared the length of residence in the United States for Asian-born and U.S.-born participants and concluded that nativity status was an important indicator of cultural exposure in our sample. Specifically, U.S.-born participants, on average, have resided in the United States 12 more years than Asian-born participants, $t(609) = 14.20$, $p < .001$. Asian-born participants have spent an average of 55.9% of their lifetime in the United States, as compared with 97.0% for U.S.-born participants.

## Measures

**Brief Symptom Inventory-18.** Participants reported their level of distress associated with 18 symptoms during the previous week on a 5-point scale ranging from 0 (*not at all*) to 4 (*extremely*). Dimension scores range from 0 to 24, and the sum of the three dimensions constitutes the global severity index (GSI), with larger numbers indicating more distress. Acceptable-to-good internal consistency was reported for the three subscales as well as the GSI in a community sample, with Cronbach's αs of .74, .84, .79, and .89 for Somatization, Depression, Anxiety, and the GSI, respectively (Derogatis, 2000). Internal consistency coefficients of the subscales and the total scale in the current study were good in all subgroups, ranging from .80 to .93.

**Rosenberg Self-Esteem Scale.** Participants rated their overall sense of worthiness as a person using the 10-item Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1979) on a 4-point rating scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). The RSES contains an equal number of positively (e.g., I feel that I have a number of good qualities) and negatively (e.g., I feel I do not have much to be proud of) worded items. In this study, negatively worded items were recoded and an average score was computed for each participant, with a higher score representing higher global self-esteem. Psychometric studies with diverse samples in terms of age, gender, and racial groups have generally supported the unidimensionality of the RSES, with some suggesting a method effect (Corwyn, 2000; Dunbar, Ford, Hunt, & Der, 2000; Thomas & Oliver, 1999). Schmitt and Allik (2005) translated and validated the RSES across 53 nations using a mixture of college student and community convenience samples. Results replicated the one-factor structure across the nations, except for Tanzania, with sufficient internal reliability ($M = .81$) and criterion-related evidence (i.e., equivalent relationships with personality factors). The RSES yielded reliable scores in the current study, with the Cronbach αs ranging from .82 to .87 across subgroups ($M = .85$).

**Racial and Life Experiences Scale.** The Racial and Life Experiences Scale (Harrell, 1997) is a battery of instruments developed to measure multiple dimensions of racism and racial socialization experiences. For this study, participants reported their racial microaggression experiences in terms of frequency and stressfulness using the 20-item Daily Life Experience subscale of the Racial and Life Experiences Scale (RaLES-DLE; e.g., being ignored, overlooked, or not given service) from 0 (*never/never happened to me*) to 5 (*once a week or more/bothers me extremely*). RaLES-DLE has shown sound psychometric properties in racially diverse samples (Harrell, 1997) and Asian American samples (Liang, Alvarez, Juang, & Liang, 2007), with the Cronbach αs ranging from .89 to .94, and theory-consistent correlations with collective self-esteem, cultural mistrust, and salience of racial identity (Harrell, 1997). In the current study, the RaLES-DLE yielded reliable scores across subgroups (.93–.97, $M = .95$). Two average scores for each participant were computed, with higher scores indicating higher frequency/stressfulness of racism experiences.

## Procedure

Participants were solicited through presentations at community and social service agencies as well as through flyers posted in local neighborhoods and commercial districts. Individuals who were interested in participating were asked to either e-mail or phone the project staff for a screening interview. Individuals were considered eligible to participate if they (a) self-identified as Vietnamese/Filipinx/Chinese or Vietnamese/Filipinx/Chinese Americans, (b) were 25 years or older, (c) lived in the United States for 10 years or more, and (d) were comfortable with speaking and reading English. Of the individuals who contacted the project staff for screening, 92% were eligible for the study, and of the eligible individuals, 95% completed the surveys. Eligible participants were then invited to schedule an appointment to have the surveys administered. Participants completed each survey packet during a 30- to 60-min face-to-face administration with the second author or his assistants. Participants were given a verbal and written debriefing as well as a $25 cash incentive.

# Results

## Test of Competing Models

Because data were collected in face-to-face administration, the data set had minimal missingness. Following recommended practices (Schlomer, Bauman, & Card, 2010), the pattern of missing data was examined. Little's missing completely at random test indicated that the data were missing completely at random. Missing data were subsequently computed using the expectation maximization strategy in IBM SPSS 20.

We aimed to identify a factor structure that fits well for all ethnicity, nativity, and gender groups. To examine the factor structure, each hypothesis was tested via confirmatory factor analysis with maximum likelihood estimation using M*plus* 6.12 (Muthén & Muthén, 1998–2011). Given the violation of multivariate normality, we used the Satorra–Bentler scaling method (Satorra & Bentler, 2001). Standardized root mean square residual values close to .09 or below, root mean square error of approximation values close to .08 or below, comparative fit index (CFI) values close to .90 or greater, and Tucker–Lewis index values exceeding .90 were considered indicative of adequate model fit (Hu & Bentler, 1999). We used the corrected/scaled likelihood ratio testing ($\Delta\chi^2$) to compare model fit of nested models (i.e., three-factor vs. one-factor) and Akaike information criterion and Bayesian information criterion to compare model fit of non-nested models (i.e., three-factor vs. bifactor).

As presented in Table 2, the one-factor model, in which the first item loading was fixed to one, showed poor fit for all ethnic, gender, and nativity groups. The three-factor model, in which the first item loading for each factor was fixed to one, exhibited significantly better fit than the one-factor model, as evidenced by likelihood ratio testing, but the model fit was only good for the female and the Asian-born groups by judging fit indices. Furthermore, we determined the three-factor model to be untenable due to moderate-to-large factor relationships that emerged for all groups between DEP and SOM (range = [.51, .74], $M = .64$, $SD = .08$), ANX and DEP (range = [.82, .91], $M = .87$, $SD = .04$), and ANX and SOM (range = [.74, .92], $M = .82$, $SD = .07$). This finding was consistent with previous BSI-18 studies with racial minority populations (Torres et al., 2013) and suggests that the three correlated factors did not represent conceptually distinct constructs for our Asian American sample.

Given the large correlations between the three factors, we tested a bifactor model in which item covariation was modeled as a general factor (GSI) and three orthogonal group factors (ANX, DEP, and SOM). We fixed the factor variances to be one for identification purposes and allowed all factor loadings to be freely estimated. The bifactor model initially failed to converge for the U.S.-born sample because the residual variance of Item 8 was estimated to be negative. We subsequently fixed this residual variance to zero for the U.S.-born sample. The bifactor model showed good fit for all groups and exhibited better fit than the three-factor model, as evidenced by smaller Akaike information criterion and Bayesian information criterion values (Table 2).

To examine the dimensionality of the bifactor model, we computed construct replicability (*H*) and explained common variance (ECV) of the general and group factors (Rodriguez, Reise, & Haviland, 2016). *H* evaluates the quality of the measurement model by a particular set of items (Hancock & Mueller, 2001). When *H* is low (<.70), the latent factor is not well defined by the items and is expected to change across studies. When *H* is high (>.70), the latent factor is well defined by the items, and in turn will have more stability across studies. *H* for the general factor ranged from .92 to .93 ($M = .92$) for all ethnicity, nativity, and gender samples. *H* for the SOM, DEP, and ANX group factors consistently fell below the recommended criterion, and ranged from .41 to .72 ($M = .58$), .39 to .62 ($M = .51$), and .23 to .63 ($M = .36$) across samples, respectively. These results indicated that the general factor likely represented a stable underlying construct, whereas the group factors were not well defined by their respective items. ECV evaluates the relative strength of the general and group factors. ECV of the general factor ranged from .68 to .79 ($M = .75$) across all samples, which suggests that the general factor explained around 68% to 79% of the common variance extracted. In conclusion, the BSI-18 bifactor model yielded strong evidence of a general factor. We retained the bifactor model for invariance testing because of its conceptual representation and

Table 2
*Fit Statistics for Brief Symptom Inventory-18 Measurement Model Across Ethnicity, Nativity, and Gender Groups*

| Model | SB $\chi^2$ | *p* | *df* | SRMR | RMSEA | CFI | AIC | BIC |
|---|---|---|---|---|---|---|---|---|
| Chinese ($n$ = 219) | | | | | | | | |
| One-factor | 345.910 | <.001 | 135 | .070 | .084 (.074,.096) | .823 | 9,027.532 | 9,210.542 |
| Three-factor | 280.193 | <.001 | 132 | .063 | .072 (.060,.083) | .876 | 8,930.201 | 9,123.378 |
| Bifactor | 221.572 | <.001 | 117 | .056 | .064 (.051,.077) | .912 | 8,873.469 | 9,117.317 |
| Filipinx ($n$ = 197) | | | | | | | | |
| One-factor | 495.848 | <.001 | 135 | .089 | .116 (.106,.128) | .719 | 8,894.492 | 9,071.785 |
| Three-factor | 317.157 | <.001 | 132 | .072 | .084 (.073,.096) | .856 | 8,646.752 | 8,833.894 |
| Bifactor | 235.487 | <.001 | 117 | .060 | .072 (.058,.085) | .908 | 8,556.932 | 8,793.323 |
| Vietnamese ($n$ = 209) | | | | | | | | |
| One-factor | 359.901 | <.001 | 135 | .074 | .089 (.078,.101) | .817 | 9,532.898 | 9,713.384 |
| Three-factor | 272.965 | <.001 | 132 | .074 | .071 (.059,.083) | .885 | 9,406.821 | 9,597.334 |
| Bi-factor | 216.813 | <.001 | 117 | .056 | .064 (.050,.077) | .919 | 9,348.434 | 9,589.082 |
| U.S.-born ($n$ = 215) | | | | | | | | |
| One-factor | 494.003 | <.001 | 135 | .089 | .111 (.101,.122) | .742 | 9,220.597 | 9,402.611 |
| Three-factor | 308.087 | <.001 | 132 | .064 | .079 (.067,.090) | .873 | 8,950.875 | 9,143.001 |
| Bifactor | 233.996 | <.001 | 118 | .066 | .068 (.055,.080) | .917 | 8,862.172 | 9,101.487 |
| Asian-born ($n$ = 405) | | | | | | | | |
| One-factor | 493.831 | <.001 | 135 | .063 | .081 (.073,.089) | .835 | 18,005.746 | 18,221.956 |
| Three-factor | 348.263 | <.001 | 132 | .060 | .064 (.056,.072) | .901 | 17,783.960 | 18,012.182 |
| Bifactor[a] | 257.555 | <.001 | 117 | .047 | .054 (.045,.063) | .935 | 17,674.642 | 17,962.922 |
| Female ($n$ = 327) | | | | | | | | |
| One-factor | 473.127 | <.001 | 135 | .066 | .088 (.079,.096) | .822 | 14,884.335 | 15,088.993 |
| Three-factor | 315.389 | <.001 | 132 | .060 | .065 (.056,.074) | .903 | 14,650.751 | 14,866.779 |
| Bifactor | 238.153 | <.001 | 117 | .047 | .056 (.046,.066) | .936 | 14,566.469 | 14,839.346 |
| Male ($n$ = 297) | | | | | | | | |
| One-factor | 511.834 | <.001 | 135 | .080 | .097 (.088,.106) | .766 | 12,524.752 | 12,724.213 |
| Three-factor | 346.249 | <.001 | 132 | .062 | .074 (.064,.083) | .867 | 12,278.646 | 12,489.189 |
| Bifactor | 273.410 | <.001 | 117 | .056 | .067 (.057,.077) | .903 | 12,193.207 | 12,459.156 |

*Note.* SB $\chi^2$ = Satorra–Bentler scaled chi-square; *df* = degrees of freedom; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; CFI = comparative fit index; AIC = Akaike information criterion; BIC = Bayesian information criterion. RMSEA values in parentheses represent 90% confidence interval.
[a] Residual variance of Item BSI8 fixed to zero for model convergence.

statistically superior model fit but refrained from interpreting results specific to the group factors (e.g., latent mean comparisons of group factors).

## Invariance Testing Strategy

Following recommended best practices (Vandenberg & Lance, 2000), a series of increasingly stringent model comparison steps were used to assess configural (i.e., same pattern of fixed and free factor loadings across groups without equality constraints), metric (i.e., equality constraints on factor loadings across groups), and (i.e., equality constraints on factor loadings and item intercepts across groups) scalar invariance. It is important to establish evidence of good model fit for each group separately (i.e., configural invariance) before conducting multigroup confirmatory factor analyses so that we can differentiate sources of model misspecification stemming from a single group versus applying equality constraint across groups. The baseline model used in metric and scalar testing simultaneously estimates model parameters for all samples under testing without any equality constraints. Likelihood ratio tests assessed whether constraining specified model parameters (e.g., factor loadings) across groups resulted in a significant worsening of model fit as compared with the baseline model. A significant result suggests a significant worsening of model fit due to equality constraints and, as a result, noninvariance. Given the number of tests required for invariance testing, a *p* value of .01 was selected a priori to reduce the probability of experimentwise error (French & Finch, 2006). Because the likelihood ratio tests are sensitive to sample size and violation of the normality assumption as chi-square tests, we also consulted changes in CFI ($\Delta$CFI $\leq -.01$ indicates invariance and $\Delta$CFI $\geq -.02$ indicates definite noninvariance) to investigate measurement invariance as a supplement (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000). It is recommended that a lack of measurement invariance be concluded when likelihood ratio tests and changes in fit indices both indicate noninvariance. When changes in fit indices indicate invariance and sample sizes are greater than 200, invariance can likely be concluded even when the likelihood ratio test is significant (Meade, Johnson, & Braddy, 2008).

**Step 1: Configural invariance.**

***Ethnicity.*** The bifactor model was individually fit to the Chinese, Filipinx, and Vietnamese samples. The bifactor model exhibited good fit for all ethnicity samples by judging fit indices (Table 2). In all three samples, all general factor item loadings were significant, and the majority were larger than the group factor item loadings (Table 3). At least half of the item loadings were nonsignificant or significantly negative for the ANX and DEP group factors in the Chinese and Filipinx samples and for the ANX group factor in the Vietnamese sample.

Table 3
*Baseline Bifactor Model Standardized Factor Loadings for Ethnicity Groups*

| | Chinese | | | | Filipinx | | | | Vietnamese | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Item | GSI | DEP | SOM | ANX | GSI | DEP | SOM | ANX | GSI | DEP | SOM | ANX |
| BSI1 | **.49** | | **.24** | | **.50** | | **.48** | | **.49** | | **.21** | |
| BSI2 | **.67** | .08 | | | **.67** | **.17** | | | **.59** | **.46** | | |
| BSI3 | **.58** | | | **−.26** | **.58** | | | −.01 | **.51** | | | .06 |
| BSI4 | **.54** | | **.25** | | **.36** | | **.46** | | **.48** | | **.39** | |
| BSI5 | **.71** | **.52** | | | **.72** | **.44** | | | **.66** | **.57** | | |
| BSI6 | **.67** | | | **−.29** | **.65** | | | −.07 | **.74** | | | −.12 |
| BSI7 | **.61** | | **.18** | | **.42** | | **.31** | | **.59** | | .15 | |
| BSI8 | **.70** | **.40** | | | **.77** | **.71** | | | **.61** | **.65** | | |
| BSI9 | **.66** | | | **.50** | **.61** | | | **.51** | **.63** | | | **.36** |
| BSI10 | **.52** | | **.26** | | **.44** | | **.56** | | **.69** | | .10 | |
| BSI11 | **.56** | **.18** | | | **.73** | −.02 | | | **.65** | **.20** | | |
| BSI12 | **.69** | | | **.28** | **.58** | | | **.45** | **.68** | | | −.10 |
| BSI13 | **.57** | | **.70** | | **.43** | | **.68** | | **.68** | | **.42** | |
| BSI14 | **.60** | **.20** | | | **.73** | −.04 | | | **.68** | **.28** | | |
| BSI15 | **.61** | | | **.23** | **.56** | | | −.10 | **.63** | | | **−.25** |
| BSI16 | **.66** | | **.47** | | **.58** | | **.51** | | **.67** | | **.43** | |
| BSI17 | **.54** | .04 | | | **.60** | −.07 | | | **.51** | **−.17** | | |
| BSI18 | **.76** | | | .06 | **.74** | | | **.31** | **.75** | | | **.24** |

*Note.* Bolded loadings significant at $p < .05$. GSI = global severity index general factor; DEP = depression group factor; SOM = somatization group factor; ANX = anxiety group factor.

Coefficient omega (ω) is a model-based reliability estimate assessing the proportion of variance attributable to all sources of common variance. Omega hierarchical ($\omega_H$) estimates the proportion of variance in total scores due to a single general factor when data are represented by a bifactor structure. A comparison between omega hierarchical and omega ($\omega_H/\omega$) represents the proportion of the reliable variance in total scores attributable to the general factor (Rodriguez et al., 2016). The general factor accounted for 97.2%, 96.9%, and 92.9% of the reliable variance in total scores for Chinese, Vietnamese, and Filipinx samples, respectively.

***Nativity.*** The bifactor model exhibited good model fit for Asian-born and U.S.-born groups by judging fit indices (Table 2). All general factor item loadings were significant, and most were larger than the group factor item loadings (Table 4). At least half of the item loadings were nonsignificant or significantly negative for the ANX group factor in the Asian-born sample and for the ANX and DEP group factors in the U.S.-born sample. The general factor accounted for 97.5.9% and 91.3% of the reliable variance in total scores for Asian-born and U.S.-born samples, respectively.

Table 4
*Baseline Bifactor Model Standardized Factor Loadings for Nativity and Gender Groups*

| | Female | | | | Male | | | | Asian-born | | | | U.S.-born | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Item | GSI | DEP | SOM | ANX | GSI | DEP | SOM | ANX | GSI | DEP | SOM | ANX | GSI | DEP | SOM | ANX |
| BSI1 | **.51** | | **.34** | | **.43** | | **.39** | | **.52** | | **.22** | | **.38** | | **.59** | |
| BSI2 | **.62** | **.31** | | | **.64** | **.26** | | | **.64** | **.32** | | | **.67** | .10 | | |
| BSI3 | **.56** | | | **.17** | **.54** | | | .01 | **.54** | | | **−.24** | **.55** | | | **.17** |
| BSI4 | **.51** | | **.42** | | **.39** | | **.38** | | **.48** | | **.32** | | **.40** | | **.52** | |
| BSI5 | **.68** | **.58** | | | **.69** | **.47** | | | **.68** | **.56** | | | **.77** | **.33** | | |
| BSI6 | **.67** | | | **.37** | **.70** | | | **−.20** | **.70** | | | −.10 | **.63** | | | .02 |
| BSI7 | **.58** | | **.19** | | **.45** | | **.29** | | **.57** | | **.17** | | **.43** | | **.40** | |
| BSI8 | **.66** | **.63** | | | **.69** | **.61** | | | **.66** | **.60** | | | **.80** | **.60** | | |
| BSI9 | **.67** | | | **−.37** | **.63** | | | **.28** | **.62** | | | .11 | **.57** | | | **.73** |
| BSI10 | **.61** | | **.27** | | **.52** | | **.38** | | **.59** | | **.23** | | **.45** | | **.58** | |
| BSI11 | **.60** | **.22** | | | **.69** | .03 | | | **.61** | **.13** | | | **.78** | −.04 | | |
| BSI12 | **.66** | | | **−.22** | **.69** | | | .01 | **.68** | | | **.33** | **.53** | | | **.53** |
| BSI13 | **.62** | | **.49** | | **.49** | | **.70** | | **.64** | | **.53** | | **.41** | | **.63** | |
| BSI14 | **.64** | **.18** | | | **.71** | .08 | | | **.67** | **.13** | | | **.76** | −.05 | | |
| BSI15 | **.62** | | | −.01 | **.57** | | | **−.21** | **.59** | | | **.32** | **.62** | | | −.02 |
| BSI16 | **.67** | | **.42** | | **.58** | | **.53** | | **.66** | | **.45** | | **.61** | | **.50** | |
| BSI17 | **.56** | −.03 | | | **.48** | −.06 | | | **.52** | −.10 | | | **.61** | −.09 | | |
| BSI18 | **.78** | | | −.03 | **.75** | | | **.52** | **.75** | | | −.04 | **.73** | | | **.35** |

*Note.* Bolded loadings significant at $p < .05$. GSI = global severity index general factor; DEP = depression group factor; SOM = somatization group factor; ANX = anxiety group factor.

***Gender.*** The bifactor model exhibited good model fit for female and male groups by judging fit indices (Table 2). All general factor item loadings were significant, and the majority were larger than the group factor item loadings (Table 4). At least half of the item loadings were nonsignificant or significantly negative for the ANX and DEP group factors in the male sample and for the ANX group factor in the female sample. The general factor accounted for 96.9% and 94.9% of the reliable variance in total scores for female and male samples, respectively.

***Summary.*** Results suggest that the bifactor model well represented the covariance relationships in the observed scores for all groups. However, raw total scores can be interpreted as an essentially unidimensional reflection of psychological distress despite the presence of multidimensionality of the data.

**Step 2: Metric invariance.**

***Ethnicity.*** The baseline model, which simultaneously estimated all model parameters in all samples without any constraints, was used for metric invariance tests. Constraining all factor loadings to be invariant across ethnic samples resulted in a nonsignificant change in model fit as compared with the baseline model, $\Delta\chi^2(72, N = 625) = 68.530$, $p = .594$, $\Delta\text{CFI} = -.008$.

***Nativity.*** Although the likelihood ratio test was significant, $\Delta\chi^2(36, N = 620) = 64.901$, $p = .002$, the change in model fit was not significant when judging the CFI, $\Delta\text{CFI} = -.008$. We concluded that constraining all factor loadings to be invariant across nativity samples resulted in a nonsignificant change in model fit as compared with the baseline model.

***Gender.*** Constraining all factor loadings to be invariant across women and men resulted in a nonsignificant change in model fit as compared with the baseline model, $\Delta\chi^2(36, N = 624) = 37.546$, $p = .398$, $\Delta\text{CFI} = -.001$. We concluded that constraining all factor loadings to be invariant across gender samples resulted in a nonsignificant change in model fit as compared with the baseline model.

***Summary.*** Results support metric invariance of the BSI-18 for the ethnicity, nativity, and gender samples. This suggests that BSI-18 items were interpreted in a similar manner by male and female, US-born and Asian born adults of Chinese, Filipinx, and Vietnamese descent.

**Step 3: Scalar invariance.**

***Ethnicity.*** Scalar invariance was examined by comparing the intercept-invariant model to the same baseline model. Constraining all item intercepts to be invariant across ethnicity samples resulted in a nonsignificant change of model fit, $\Delta\chi^2(108, N = 625) = 130.036$, $p = .073$, $\Delta\text{CFI} = -.017$.

***Nativity.*** Constraining all intercept terms to be invariant across nativity samples resulted in a significant worsening of model fit as compared with the baseline model, $\Delta\chi^2(54, N = 620) = 120.020$, $p < .001$, $\Delta\text{CFI} = -.018$. The modification index (MI) provides an estimated value in which the model's chi-square test would decrease if a parameter were added and freely estimated in the model. MI greater than 3.841 indicates that misspecification of the corresponding parameters significantly contributes to a lack of model fit. Although MI for other parameters (e.g., uniqueness term covariance) was much larger than MI for the intercepts, we only considered freeing intercept equality constraints for theoretical clarity and meaningfulness. However, because we did not have theoretical reasons to explain specific intercepts to be different across Asian-born and U.S.-born groups, we relied on the magnitude of MI of the intercepts to determine the sequence of model modification. Intercepts for Items 6 (MI = 8.636), 15 (MI = 5.056), 1 (MI = 4.462), and 4 (MI = 4.852) were noninvariant across Asian-born and U.S.-born samples, and the equality constraint on these intercepts was in turn sequentially freed. The intercept partially invariant model exhibited a nonsignificant difference in fit as compared with the baseline model, $\Delta\chi^2(50, N = 620) = 92.844$, $p < .001$, $\Delta\text{CFI} = -.012$.

***Gender.*** Constraining all item intercepts to be invariant across gender samples resulted in a nonsignificant change in model fit as compared with the baseline model, $\Delta\chi^2(54, N = 624) = 68.156$, $p = .093$, $\Delta\text{CFI} = -.006$. We concluded that constraining all item intercepts to be invariant across gender samples resulted in a nonsignificant change in model fit as compared with the baseline model.

***Summary.*** Results support at least partial scalar invariance of the BSI-18 for the ethnicity, nativity, and gender samples. Because scalar invariance was a requirement for latent mean identification, we proceeded to testing latent mean invariance.

**Step 4: Invariance of latent means.** Latent mean analysis is the latent equivalence of traditional ANOVA and is superior because it is not biased by measurement error. To test whether the latent mean of the BSI-18 factors was invariant across ethnicity, nativity, and gender samples, one group (e.g., the Vietnamese group) was arbitrarily set to be the reference group. The factor means of the reference group were set to be zero, whereas the latent means of other groups were freely estimated. The estimated latent means represented the difference between the reference group and the other groups, where significant positive values indicated higher levels of the latent factor than the reference group, significant negative values indicated lower levels of the latent factor than the reference group, and nonsignificant values indicated latent mean invariance between groups. The Chinese group and the Filipinx group did not differ from the Vietnamese group (reference group) on mean levels of the general factor, $\hat{\xi}_{GSI} = -.080$ and .052, $p = .436$ and .640, respectively. Women had higher mean levels of the general factor than men (reference group) by .302, $p = .001$. U.S.-born and Asian-born (reference group) Asian Americans, on average, did not differ on levels of the general factor, $\hat{\xi}_{GSI} = -.079$, $p = .416$.

## Criterion-Related Evidence

The ancillary bifactor measures indicated that the BSI-18 yielded a strong general factor in our sample, both in terms of its raw scores and in the structural equation modeling context. As a result, we tested criterion-related evidence using total raw scores of the BSI-18 by summing all items. Consistent with our hypotheses, the BSI-18 total score was moderately and negatively correlated with self-esteem in all gender, nativity, and ethnicity subsamples, $r$s = [−.39, −.60], $p$s < .001, which suggests that participants with higher levels of self-esteem tend to have lower psychological distress as measured by BSI-18. In addition, the BSI-18 total score was positively associated with racial microaggression in terms of frequency and perceived stress in all gender, nativity, and ethnicity subsamples, $r$s = [.25, .41], $p$s < .001, which suggests that participants who reported more stressful and frequent racial microaggression experiences tend to have more psychological distress as measured by the BSI-18.

## Discussion

This is the first study to test the factor structure and measurement invariance of the BSI-18 across ethnicity (Vietnamese, Filipinx, and Chinese), nativity status (U.S.-born vs. Asian-born), and gender groups (women vs. men) of Asian Americans. Similar to previous studies with other racial minority individuals (Torres et al., 2013), the multidimensionality of the BSI-18 was called into question due to large factor relationships in the three-factor model. Although the bifactor model was retained for conceptual and statistical superiority, ancillary measures suggest that the general factor had strong evidence and explained the majority proportion of raw score variance and extracted common variance, whereas the group factors needed further examination. We thus recommend that for adults of Asian ethnic descent, the BSI-18 be used as an instrument of general psychological distress rather than as subscale scores.

We found evidence of configural and metric invariance of the BSI-18 across ethnicity, nativity, and gender when modeled as a bifactor structure, which suggests that items were interpreted in a similar manner by male and female, U.S.-born and Asian-born adults of Chinese, Filipinx, and Vietnamese descent. Scalar invariance of the BSI-18 was evidenced when assessing item intercepts across ethnicity and gender, and partial scalar invariance was established for nativity groups. Specifically, intercepts for two anxiety items (i.e., "feeling tense or keyed up" and "feeling so restless you couldn't sit still") and two somatic items (i.e., "faintness or dizziness" and "pains in heart or chest") were noninvariant for nativity groups. The two anxiety item intercepts were estimated to be higher for the U.S.-born group, and the two somatic item intercepts were estimated to be higher for the Asian-born group. Although it is unclear whether the intercept noninvariance across nativity samples was a systemic pattern or a sample-specific finding, it seems that the two anxiety items overestimated distress for U.S.-born individuals, whereas the somatic items overestimated distress for Asian-born individuals. It is possible that acculturation is related to symptom expression for Asian Americans, where Asian-born individuals endorse more somatic symptoms (i.e., dizziness and heartache) and U.S.-born individuals endorse more anxiety symptoms (i.e., tenseness and restless). However, most item intercepts were invariant for nativity samples. We conclude that BSI-18 item mean scores are overall comparable across ethnicity, gender, and nativity groups.

At the structural level, women, on average, endorsed higher levels of the latent GSI factor than men. The relationship between gender and psychological distress among Asian Americans is inconclusive in the literature (Chung & Kagawa-Singer, 1993; Lee et al., 2015). This is understandable given the complicated relationship between gender and other internal and external processes, such as coping (Liang et al., 2007) and experiences of discrimination (Nadal, Wong, Sriken, Griffin, & Fujii-Doe, 2015), which all have a bearing on mental health. In addition, we found that participants of Chinese, Filipinx, and Vietnamese descent reported similar levels of the latent factor. Researchers typically theorize potential differences in mental health between Asian American ethnic groups largely stem from different preimmigration (e.g., trauma and preimmigration SES) and postimmigration (e.g., access to employment and racism experiences; Takeuchi et al., 2007) experiences. The nonsignificant latent mean differences found in our study added confidence of using the BSI-18 in Asian ethnic group comparisons but could not disentangle the complex factors that may contribute to ethnic group-level differences or a lack thereof. Finally, we found that Asian-born and U.S.-born groups reported similar mean levels of the latent factor. Past studies on the mental health of immigrants found that foreign-born Asians and Latinx report better mental health than their U.S.-born counterparts (Alegría et al., 2007; Breslau & Chang, 2006). Alegría et al. (2008) first coined the term *immigrant paradox* to describe the conflicting findings that construe immigrant status as both a risk and protective factor for Asian American and Latinx populations. Leong, Park, and Kalibatseva (2013) attempted to disentangle the complex relationship between immigrant status and mental health by identifying both protective (e.g., social networking) and risk (e.g., acculturative stress and family conflict) factors associated with immigrant status. As a result, the nonsignificant latent mean comparison between U.S.-born and Asian-born groups should be interpreted with caution, as it may be due to the manifestation of interacting protective and risk factors in association with immigration status.

Findings suggest external validity of the BSI-18 across gender, nativity, and ethnic subsamples, as evidenced by hypothesis-consistent correlations with self-esteem and racial microaggression experiences. Ultimately, results provide evidence for the calculation and interpretation of the total scores of the BSI-18 as a measure of general psychological distress among adults of Asian ethnic descent in the community across nativity status and gender. We recommend against the use of subscale scores or the interpretation of the group factors in a structural equation modeling context.

Present findings should be evaluated in light of limitations. First, although the study participants included Chinese, Filipinx, and Vietnamese American adults, they did not represent the breadth of diversity of the Asian American population. It is not clear whether and how these findings would apply across other Asian ethnic groups who have a unique sociopolitical experience relative to the current sample, such as South Asians or Asian elders. Second, the joint associations of nativity and gender were not examined in the study, although mental health in reality is impacted by experiences stemming from intersecting identities (Hong et al., 2014). Third, our findings are based on a community adult sample. It is unclear whether present findings would replicate across Asian American populations with clinical and/or medical disorders, or across populations from varying SES backgrounds, as these factors may also influence symptom presentation. Finally, although competing models of the BSI-18 were tested, there are likely plausible alternative models that were not tested and therefore cannot be ruled out. Researchers are strongly encouraged to conduct further internal structure analyses of the BSI-18 in Asian American samples to determine the idiosyncrasy versus generalizability of our findings.

## References

Alegría, M., Canino, G., Shrout, P. E., Woo, M., Duan, N., Vila, D., . . . Meng, X.-L. (2008). Prevalence of mental illness in immigrant and non-immigrant U.S. Latino groups. *The American Journal of Psychiatry, 165,* 359–369. http://dx.doi.org/10.1176/appi.ajp.2007.07040704

Alegría, M., Mulvaney-Day, N., Torres, M., Polo, A., Cao, Z., & Canino, G. (2007). Prevalence of psychiatric disorders across Latino subgroups in the United States. *American Journal of Public Health, 97,* 68–75. http://dx.doi.org/10.2105/AJPH.2006.087205

Alvarez, A. N., Juang, L., & Liang, C. T. (2006). Asian Americans and racism: When bad things happen to "model minorities". *Cultural Diver-*

*sity and Ethnic Minority Psychology, 12,* 477–492. http://dx.doi.org/10.1037/1099-9809.12.3.477

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York, NY: Harber Medical Division, Harper and Row.

Breslau, J., & Chang, D. F. (2006). Psychiatric disorders among foreign-born and U. S.-born Asian-Americans in a U.S. national survey. *Social Psychiatry and Psychiatric Epidemiology, 41,* 943–950. http://dx.doi.org/10.1007/s00127-006-0119-2

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5

Chung, R. C.-Y., & Kagawa-Singer, M. (1993). Predictors of psychological distress among Southeast Asian refugees. *Social Science and Medicine, 36,* 631–639. http://dx.doi.org/10.1016/0277-9536(93)90060-H

Clark, R., Anderson, N. B., Clark, V. R., & Williams, D. R. (1999). Racism as a stressor for African Americans: A biopsychosocial model. *American Psychologist, 54,* 805–816. http://dx.doi.org/10.1037/0003-066X.54.10.805

Corwyn, R. F. (2000). The factor structure of global self-esteem among adolescents and adults. *Journal of Research in Personality, 34,* 357–379. http://dx.doi.org/10.1006/jrpe.2000.2291

Derogatis, L. R. (2000). *The Brief Symptom Inventory-18 (BSI-18): Administration, scoring, and procedures manual* (3rd ed.). Minneapolis, MN: National Computer Systems.

Dunbar, M., Ford, G., Hunt, K., & Der, G. (2000). Question wording effects in the assessment of global self-esteem. *European Journal of Psychological Assessment, 16,* 13–19. http://dx.doi.org/10.1027//1015-5759.16.1.13

French, B. F., & Finch, W. H. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling, 13,* 378–402. http://dx.doi.org/10.1207/s15328007sem1303_3

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural Equation Modeling: Rresent and Future. A festschrift in honor of Karl Jöreskog* (pp. 195–216). Lincolnwood, IL: Scientific Software International.

Harrell, S. P. (1997). *Development and initial validation of scales to measure racism-related stress*. In *Proceedings of the Sixth Biennial Conference on Community Research and Action*. Columbia, SC: Society for Community Research and Action.

Hong, S., Walton, E., Tamaki, E., & Sabin, J. A. (2014). Lifetime prevalence of mental disorders among Asian Americans: Nativity, gender, and sociodemographic correlates. *Asian American Journal of Psychology, 5,* 353–363. http://dx.doi.org/10.1037/a0035680

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Iwamasa, G. Y., & Hilliard, K. M. (1999). Depression and anxiety among Asian American elders: A review of the literature. *Clinical Psychology Review, 19,* 343–357. http://dx.doi.org/10.1016/S0272-7358(98)00043-9

Joiner, T. E., Katz, J., & Lew, A. (1999). Harbingers of depressotypic reassurance seeking: Negative life events, increased anxiety, and decreased self-esteem. *Personality and Social Psychology Bulletin, 25,* 632–639. http://dx.doi.org/10.1177/0146167299025005008

Kalibatseva, Z., Leong, F. T. L., Ham, E. H., Lannert, B. K., & Chen, Y. (2017). Loss of face, intergenerational family conflict, and depression among Asian American and European American college students. *Asian American Journal of Psychology, 8,* 126–133. http://dx.doi.org/10.1037/aap0000067

Kearney, L. K., Draper, M., & Barón, A. (2005). Counseling utilization by ethnic minority college students. *Cultural Diversity and Ethnic Minority Psychology, 11,* 272–285. http://dx.doi.org/10.1037/1099-9809.11.3.272

Kernis, M. H., Grannemann, B. D., & Mathis, L. C. (1991). Stability of self-esteem as a moderator of the relation between level of self-esteem and depression. *Journal of Personality and Social Psychology, 61,* 80–84. http://dx.doi.org/10.1037/0022-3514.61.1.80

Kim, B. S., Yang, P. H., Atkinson, D. R., Wolfe, M. M., & Hong, S. (2001). Cultural value similarities and differences among Asian American ethnic groups. *Cultural Diversity and Ethnic Minority Psychology, 7,* 343–361. http://dx.doi.org/10.1037/1099-9809.7.4.343

Lee, M. C., Friedmann, E., Kverno, K., Newhouse, R., Zhang, D., & Thomas, S. (2015). Psychological distress among Chinese immigrants to the USA. *International Journal of Culture and Mental Health, 8,* 150–161. http://dx.doi.org/10.1080/17542863.2014.913643

Leong, F., Park, Y. S., & Kalibatseva, Z. (2013). Disentangling immigrant status in mental health: Psychological protective and risk factors among Latino and Asian American immigrants. *American Journal of Orthopsychiatry, 83* (Part 3), 361–371. http://dx.doi.org/10.1111/ajop.12020

Liang, C. T. H., Alvarez, A. N., Juang, L. P., & Liang, M. X. (2007). The role of coping in the relationship between perceived racism and racism-related stress for Asian Americans: Gender differences. *Journal of Counseling Psychology, 54,* 132–141. http://dx.doi.org/10.1037/0022-0167.54.2.132

Liu, C. M., & Suyemoto, K. L. (2016). The effects of racism-related stress on Asian Americans: Anxiety and depression among different generational statuses. *Asian American Journal of Psychology, 7,* 137–146. http://dx.doi.org/10.1037/aap0000046

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93,* 568–592. http://dx.doi.org/10.1037/0021-9010.93.3.568

Miller, M. J., Alvarez, A. N., Li, R., Chen, G. A., & Iwamoto, D. K. (2016). Measurement invariance of the people of Color Racial Identity Attitudes Scale with Asian Americans. *Psychological Assessment, 28,* 116–122. http://dx.doi.org/10.1037/pas0000161

Miller, M. J., Kim, J., & Benet-Martínez, V. (2011). Validating the Riverside Acculturation Stress Inventory with Asian Americans. *Psychological Assessment, 23,* 300–310. http://dx.doi.org/10.1037/a0021589

Miller, M. J., Yang, M., Farrell, J. A., & Lin, L. (2011). Racial and cultural factors affecting the mental health of Asian Americans. *American Journal of Orthopsychiatry, 81,* 489–497. http://dx.doi.org/10.1111/j.1939-0025.2011.01118.x

Muthén, L. K., & Muthén, B. O. (1998–2011). *Mplus user's guide* (6th ed.). Los Angeles, CA: Author.

Nadal, K. L., Wong, Y., Sriken, J., Griffin, K., Fujii-Doe, W. (2015). Racial microaggressions and Asian Americans: An exploratory study on within-group differences and mental health. *Asian American Journal of Psychology, 6,* 136–144. http://dx.doi.org/10.1037/%2Fa00380

Okazaki, S. (2009). Impact of racism on ethnic minority mental health. *Perspectives on Psychological Science, 4,* 103–107. http://dx.doi.org/10.1111/j.1745-6924.2009.01099.x

Orth, U., Robins, R. W., & Roberts, B. W. (2008). Low self-esteem prospectively predicts depression in adolescence and young adulthood. *Journal of Personality and Social Psychology, 95,* 695–708. http://dx.doi.org/10.1037/0022-3514.95.3.695

Prelow, H. M., Weaver, S. R., Swenson, R. R., & Bowman, M. A. (2005). A preliminary investigation of the validity and reliability of the Brief-Symptom Inventory-18 in economically disadvantaged Latina American mothers. *Journal of Community Psychology, 33,* 139–155. http://dx.doi.org/10.1002/jcop.20041

Recklitis, C. J., Parsons, S. K., Shih, M. C., Mertens, A., Robison, L. L., & Zeltzer, L. (2006). Factor structure of the Brief Symptom Inventory-18 in adult survivors of childhood cancer: Results from the childhood cancer survivor study. *Psychological Assessment, 18,* 22–32. http://dx.doi.org/10.1037/1040-3590.18.1.22

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21,* 137–150. http://dx.doi.org/10.1037/met0000045

Rosenberg, M. (1979). *Conceiving the self.* New York, NY: Basic Books.

Ryder, A. G., Yang, J., Zhu, X., Yao, S., Yi, J., Heine, S. J., & Bagby, R. M. (2008). The cultural shaping of depression: Somatic symptoms in China, psychological symptoms in North America? *Journal of Abnormal Psychology, 117,* 300–313. http://dx.doi.org/10.1037/0021-843X.117.2.300

Sanchez, F., & Gaw, A. (2007). Mental Health Care of Filipino Americans. *Psychiatric Services, 58,* 810–815.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66,* 507–514. http://dx.doi.org/10.1007/BF02296192

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology, 89,* 623–642. http://dx.doi.org/10.1037/0022-3514.89.4.623

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology, 57,* 1–10. http://dx.doi.org/10.1037/a0018082

Takeuchi, D. T., Zane, N., Hong, S., Chae, D. H., Gong, F., Gee, G. C., . . . Alegría, M. (2007). Immigration-related factors and mental disorders among Asian Americans. *American Journal of Public Health, 97,* 84–90. http://dx.doi.org/10.2105/AJPH.2006.088401

Thomas, J. M., & Oliver, A. (1999). Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling, 6,* 84–98. http://dx.doi.org/10.1080/10705519909540120

Torres, L., Miller, M. J., & Moore, K. M. (2013). Factorial invariance of the Brief Symptom Inventory-18 (BSI-18) for adults of Mexican descent across nativity status, language format, and gender. *Psychological Assessment, 25,* 300–305. http://dx.doi.org/10.1037/a0030436

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3,* 4–70. http://dx.doi.org/10.1177/109442810031002

Wiesner, M., Chen, V., Windle, M., Elliott, M. N., Grunbaum, J. A., Kanouse, D. E., & Schuster, M. A. (2010). Factor structure and psychometric properties of the Brief Symptom Inventory-18 in women: A MACS approach to testing for invariance across racial/ethnic groups. *Psychological Assessment, 22,* 912–922. http://dx.doi.org/10.1037/a0020704

Wong, R., Wu, R., Guo, C., Lam, J. K., & Snowden, L. R. (2012). Culturally sensitive depression assessment for Chinese American immigrants: Development of a comprehensive measure and a screening scale using an item response approach. *Asian American Journal of Psychology, 3,* 230–253. http://dx.doi.org/10.1037/a0025628

Yen, S., Robins, C. J., & Lin, N. (2000). A cross-cultural comparison of depressive symptom manifestation: China and the United States. *Journal of Consulting and Clinical Psychology, 68,* 993–999. http://dx.doi.org/10.1037/0022-006X.68.6.993

Yoo, H. C., Miller, M. J., & Yip, P. (2015). Validation of the internalization of the Model Minority Myth Measure (IM-4) and its link to academic performance and psychological adjustment among Asian American adolescents. *Cultural Diversity and Ethnic Minority Psychology, 21,* 237–246. http://dx.doi.org/10.1037/a0037648

Zane, N. (2007). *Asian American Center on Disparities Research: Community-based research on EBPPs*. Symposium conducted at the annual convention of the Asian American Psychological Association, San Francisco, CA.

Received November 6, 2017
Revision received February 20, 2018
Accepted February 20, 2018 ■