Article

Measurement Invariance of the Prodromal Questionnaire–Brief Among White, Asian, Hispanic, and Multiracial Populations

Assessment
2019, Vol. 26(2) 294–304

DOI: 10.1177/1073191111668739l
journals.sagepub.com/home/asm
SAGE

David C. Cicero[1], Alexander Krieg[1], and Elizabeth A. Martin[2]

Abstract

The Prodromal Questionnaire–Brief is a scale that is used to screen individuals for risk for the development of psychosis. It has promising psychometric properties in clinical and nonclinical populations, including undergraduates. However, the measurement invariance of the scale has not been examined in Asian, White, Hispanic, and Multiracial samples. A total of 2,767 undergraduates at two large public U.S. universities completed the Prodromal Questionnaire–Brief. The Total scores had configural and scalar invariance, while the Distress scores displayed configural, metric, and partial scalar invariance. Follow-up analyses revealed that three items were responsible for the lack of complete scalar invariance for the Distress scores. This suggests that the Total and Distress scores are measuring the same construct across groups and mean scores represent the same level of latent prodromal traits across groups. Mean comparisons for the Distress Scale across ethnicity should be interpreted with caution because it lacks complete scalar invariance. White and Hispanic participants had lower Total scores that Multiracial and Asian participants, and this pattern emerged for 13 items. For the distress items that were scalar invariant, the Asian group reported more distress than the White and Hispanic groups, while the Multiracial group reported more distress than the White group.

Keywords

prodrome, psychosis, measurement invariance, ethnicity, schizotypy, psychotic-like experiences, attenuated psychotic symptoms

The schizophrenia prodrome is a period of days, weeks, or months between the onset of attenuated psychotic symptoms and full-blown psychosis (Moller & Husby, 2000). Recent efforts have aimed to identify people in this stage of the disorder, or in the preceding premorbid phase, so that they may receive appropriate treatment (Addington & Heinssen, 2012). One factor that has been associated with poor prognosis of psychosis is duration of untreated psychosis, which refers to the period of time between the onset of a first psychotic episode and the first treatment (Marshall et al., 2005). Early identification of psychosis or psychosis-risk may reduce duration of untreated psychosis by enabling clinicians to monitor young people who are at risk for psychosis and intervene when appropriate, which may delay or potentially prevent its onset altogether (Addington & Heinssen, 2012).

The assessment of risk for psychosis is typically done with structured interviews such as the Structured Interview for Prodromal Syndromes (Miller et al., 2003) and the Comprehensive Assessment of At-Risk Mental States (Yung et al., 2005). These interviews have been shown to accurately identify individuals at risk for psychosis (Cannon et al., 2008), and early intervention programs around the world use them to determine whether potential clients are appropriate for treatment (Fusar-Poli et al., 2013). Despite these advances, many young people do not get into treatment until they have already progressed beyond the prodrome and into a first episode of psychosis (Addington, Van Mastrigt, Hutchinson, & Addington, 2002). One strategy that has been proposed to address this gap is to screen general population samples for psychosis, such as college students, and refer people with high scores for appropriate assessment and treatment services (Kline & Schiffman, 2014). However, these comprehensive interviews may be impractical with such populations because psychosis risk has a low base rate, which would require a large number of people to be interviewed to detect people at risk. Moreover, these interviews require trained interviewers to spend time

[1]University of Hawaii at Manoa, Honolulu, HI, USA
[2]University of California, Irvine, CA, USA

Corresponding Author:
David C. Cicero, Department of Psychology, University of Hawaii at Manoa, 2530 Dole Street, Sakamaki D-406, Honolulu, HI 96822-2294, USA.
Email: dcicero@hawaii.edu

with individual clients (Kline & Schiffman, 2014). Thus, researchers have developed brief screening instruments that can be given to large numbers of people in schools, colleges, or other community-based settings (Loewy, Pearson, Vinogradov, Bearden, & Cannon, 2011).

One of the most commonly used psychosis-risk screening instruments is the Prodromal Questionnaire–Brief (PQ-B; Loewy et al., 2011), a shortened version of the 92-item Prodromal Questionnaire (PQ; Loewy, Bearden, Johnson, Raine, & Cannon, 2005). The PQ-B has been shown to be highly correlated with structured interview ratings of prodromal symptoms and to have high sensitivity and specificity in "diagnosing" individuals as prodromal (i.e., having attenuated psychosis syndrome) using the SIPS as the final "diagnostic" instrument (e.g., Kline et al., 2012; Loewy et al., 2011). Previous research on people putatively in the schizophrenia prodrome has found that attenuated positive symptoms are the best predictors of "conversion" to psychosis (see Addington & Heinssen, 2012, for a review). As a result, the proposed criteria for attenuated psychosis syndrome in the *Diagnostic and Statistical Manual of Mental Disorders–5th Edition* (American Psychiatric Association, 2013) includes attenuated positive, but not negative, disorganized or general symptoms among its criteria. Since the PQ-B was developed to be a screener for the same syndrome, it also only includes positive symptoms. The PQ-B is a relatively brief instrument that only includes 21 items and gives both a "Total" score and a "Distress" score. Moreover, the PQ-B may be an especially good measure for use in college students because the original development studies included college student participants.

Examining the measurement invariance of the PQ-B in college students is important for both the provision of clinical services and for research. One use of the PQ-B could be to screen for psychosis in college students. College counseling centers are serving increasing numbers of college students for a variety of mental illnesses (Benton, Robertson, Tseng, Newton, & Benton, 2003), which has been attributed to better understanding of mental illness in college students (Hunt & Eisenberg, 2010). These counseling centers could also screen for risk, psychosis, and refer those found to be at risk for a more thorough assessment. Undergraduates are in college during the age range for the highest risk of psychosis onset (Häfner et al., 1998). As a result, college counseling centers have a higher concentration of young people in the risk period than do clinics that see only adults or only children. College students may be a convenient population to target for clinicians and researchers with limited resources. At the same time, many research studies have used the PQ and the PQ-B with undergraduates (e.g., Cooper, Klugman, Heimberg, Anglin, & Ellman, 2016; Denenny, Thompson, Pitts, Dixon, & Schiffman, 2015; DeVylder, Jahn, et al., 2015; DeVylder, Thompson, Reeves, & Schiffman, 2015; Gibson et al., 2014; Mittal, Dean, Pelletier, & Caligiuri, 2011). Understanding the measurement invariance of the PQ-B is essential for the proper interpretation of these studies. Thus, even if these results do not generalize beyond college students, examining the measurement invariance of the PQ-B in college students will provide useful information to clinicians and researchers working with college populations.

Undergraduates may have higher socioeconomic status and more education than their age-matched peers who do not attend college, but there are several reasons to suggest that results from studies with undergraduates would generalize to these populations. First, previous research suggests that attenuated psychotic symptoms are common in the general population (van Os, Hanssen, Bijl, & Vollebergh, 2001), including among undergraduates (Cicero, Martin, Becker, Docherty, & Kerns, 2014; Loewy, Johnson, & Cannon, 2007). However, it should be noted that the many of these symptoms are experienced infrequently and the more common symptoms may not be indicative of risk for psychosis (Yung et al., 2009). Second, recent research suggests that college students do not differ from their peers who do not attend college in overall levels of psychopathology (Blanco et al., 2008). In one recent study in China, a similar measure to the PQ-B, the PQ-16, was used to successfully screen college students who were then further assessed with the SIPS (F. Chen et al., 2014). Fifty-four undergraduates out of a pool of 589 exceeded the cut score on the PQ-16. Of these 54 undergraduates, 20 met criteria for attenuated psychosis syndrome. Thus, screening undergraduates shows promise in proactively identifying people at risk for psychosis.

Although the PQ-B is potentially a good candidate to serve as this screening measure, its psychometric properties have not been extensively examined in some populations such as Asian, Multiracial, and Hispanic college students. For example, cut points have been suggested for risk for psychosis, but it is unclear if these cut points can be applied broadly across ethnic groups. One way to examine whether the cut scores should be the same or different across groups is to test measurement invariance.

Examining the measurement invariance of the PQ-B in these samples is also important for several reasons. First, numerous studies have found differences in levels of psychotic-like experiences in these populations (e.g., W. J. Chen, Hsiao, & Lin, 1997; Chmielewski, Fernandes, Yee, & Miller, 1995; Cicero, 2015; Schiffman, 2004), and these differences may be related to a variety of cultural factors. For example, increased rates of psychosis and psychotic-like experiences in ethnic minorities have been linked to experiences of discrimination or racism (Anglin, Lighty, Greenspoon, & Ellman, 2014; Chakraborty & McKenzie, 2002; Karlsen, Nazroo, McKenzie, Bhui, & Weich, 2005), acculturative stress (Devylder et al., 2013), low levels in strength of ethnic identity (Burnett-Zeigler, Bohnert, &

Ilgen, 2013), and increased social adversity related to minority status, among other factors (Wicks, Hjern, Gunnell, Lewis, & Dalman, 2005). However, without testing measurement invariance of the scales, it is not possible to tell whether these differences represent actual differences in levels of prodromal symptoms or are an artifact of the psychometric properties of the measurements. Second, elevated rates of psychotic-like experiences observed in some samples may be due to cultural variations in which beliefs are culturally appropriate. If a belief is shared by an individual's culture, then it is not considered a psychotic or attenuated psychotic symptom even if it is considered abnormal by the majority culture (American Psychiatric Association, 2013). Measurement invariance tests would provide insight into whether this is occurring. If the PQ-B is shown to have metric and scalar invariance in these groups, it would suggest that cultural differences in responding are not related to different rates of psychotic-like experiences. In contrast, if the scales are not shown to have metric or scalar invariance, it would suggest that cultural differences in responding may be driving the group differences and that the group differences may not be related to real differences between groups in rates of psychotic-like experiences. Third, some research suggests that similar measures of psychotic-like experiences are not invariant across White, Multiracial, and Asian samples (Cicero, 2015) and between White and other ethnic minority samples (Winterstein, Ackerman, Silvia, & Kwapil, 2011). It is unclear if the PQ-B also lacks measurement invariance in these populations. Fourth, testing the measurement invariance of the PQ-B in these populations is a crucial first step in appropriately using the PQ-B to screen for psychosis in diverse samples. Studies that include diverse samples implicitly assume that the scores have the same meaning between groups, but this assumption can, and should, be tested empirically. To our knowledge, no studies have examined the measurement invariance of the PQ-B across any ethnic groups.

The primary goal of the current research was to examine the configural, metric, and scalar invariance of the PQ-B Total Scale and Distress Scale scores in White, Hispanic, Asian, and Multiracial undergraduates. The second goal of the current research is to examine whether there are ethnic differences in PQ-B scores, as has been found with similar measures in previous research.

Method

Participants

Participants were 3,081 undergraduates enrolled in psychology courses from a large public university in Hawaii and a large public West Coast University who participated for either partial completion of a course requirement or course extra credit. There were 1,496 Asian including 946 East Asian (Chinese, Japanese, Korean, etc.), 492 Southeast Asian (Vietnamese, Cambodian, Filipino, etc.), and 58 South Asian (Indian, Pakistani, Sri Lankan, etc.), 604 White, 419 Hispanic, and 322 Multiracial participants. Sixty-eight participants who identified as African American, 131 who identified as Native American, Native Hawaiian, or Alaska Native, and 41 participants who identified as Pacific Islander were excluded from the analyses due to small sample sizes. Thirty-seven participants older than age 35 were excluded because schizophrenia tends to have an onset in late adolescence/early adulthood (Häfner et al., 1998). An additional 68 participants were excluded listwise for having missing data, leaving a final sample of 2,738 including 1,445 Asian, 573 White, 413 Hispanic, and 307 Multiracial participants. Age ranged from 17 to 35 (M = 20.08, SD = 2.521). They were 72.3% female. The State of Hawaii sample included was 53.9% Asian, 25.5% White, 5.4% Hispanic, and 15.2% Multiracial. The West Coast sample was 51.9% Asian, 14.0% White, 28.4% Hispanic, and 5.7% Multiracial. These samples were statistically significantly different in terms of ethnicity, $\chi^2(3) = 340.305$, $p < .001$, $\varphi = .35$. The Hawaii sample was 66.1% female, while the West Coast sample was 84.4% female, $\chi^2(1) = 115.611$, $p < .001$, $\varphi = .20$. The mean age of the Hawaii sample was 19.81 (SD = 2.60), while the mean age of the West Coast sample was 20.43 (SD = 2.21), which was also statistically significant, $t(2,779) = 6.613$, $p < .001$, $d = .25$. Although the samples were different in terms of demographics, all the differences had relatively small effect sizes and were significant due to the large sample sizes. There were no significant differences in PQ-Total or Distress scores between people of the same ethnicity between the Hawaii and West Coast samples (e.g., between Asian American participants in Hawaii and Asian American participants on the West Coast).

There was an unequal distribution across ethnicities in sex of the participants. The Asian group was 69.9% female, the White group was 75.0% female, the Hispanic group was 85.3% female, and the Multiracial group was 72.3% female, which was a statistically significant difference, $\chi^2(3) = 40.171$, $p < .001$.

Materials

Participants completed the PQ-B (Loewy et al., 2011) online as part of a larger study. The PQ-B is a 21-item questionnaire that was abbreviated from the original 92-item PQ. In the development of the PQ-B, Loewy et al. (2011) retained only the positive items of the PQ because the positive items are critical items for the structured interview "diagnoses" of psychosis risk. They then removed items that were endorsed by a high percentage of undergraduates and selected the items that were the most strongly correlated with SIPS diagnoses.

Finally, they tested the items to see which best predicted SIPS diagnoses and ended with 21 total items. Participants answer each question either *yes* or *no*. For each affirmative answer, participants are instructed to answer a follow-up question: "When this happens, I feel frightened, concerned, or it causes problems for me" on a scale from 1 (*strongly disagree)* to 5 (*strongly agree)*. The final scale yields two scores: (a) the Total score, the sum of the affirmative answers (i.e., *no* = 0, *yes* = 1), and (b) a Distress score, the total number of endorsed positive symptom items weighted by level of distress (i.e., *no* = 0, *yes: strongly disagree* = 1, *disagree* = 2, *neutral* = 3, *agree* = 4, *strongly agree* = 5). In the current research, the Total Scale had an internal consistency of $\alpha = .86$ and the Distress score had an internal consistency of $\alpha = .89$.

Data Analytic Strategy

All measurement invariance analyses were conducted with M*plus* version 7.31 (Muthen & Muthen, 1998-2016). As mentioned, the PQ-B yields two scores including a Total score and a Distress score. Thus, we tested the measurement invariance of both scoring schemes. First, we fit a one-factor model for all the data in single group analysis to verify that the one-factor model fit the data well. Second, we tested the measurement invariance of the PQ-B dichotomous yes/no scores using a multiple group confirmatory factor analysis. We tested the fit of a one-factor configural model in which the factor loadings and intercepts are allowed to vary in all four ethnic groups. Next, we tested the fit of a scalar invariance model in which the factor loadings and intercepts are constrained to be equal across the groups. In all of these models, all data were specified as "categorical," thus using a polychoric as opposed to a Pearson correlation matrix. We used a weighted least squares, mean and variance (WLSMV) adjusted estimation method, as recommended for categorical data (Brown, 2006; Muthen & Muthen, 1998-2016). We did not test the fit of a metric invariance model (i.e., constraining the loadings to be equal among groups, but not the intercepts) because this is not appropriate for a scale composed of only dichotomous variables (Muthen & Muthen, 1998-2016).

Next, we tested the measurement invariance of the Distress Scale. Like the Total Scale score, we first specified a one-factor model with WLSMV for a single group with all the data to confirm that a one-factor model fit the data well. WLSMV requires that participants from each group select each response option for multiple group confirmatory factor analyses. In the current study, no Hispanic participants selected "5" for one item and no Multiracial participants selected "5" for two items. Thus, we tested the measurement invariance models with maximum likelihood estimation method (MLM). We specified a one-factor configural model in which the factor loadings and intercepts were free to vary among all four ethnic groups. Then, we tested a metric invariance model in which the loadings were constrained to be equal across groups. Finally, we tested a scalar invariance model in which the loadings and intercepts were constrained to be equal among the four groups.

To determine whether the scales have measurement variance, the fit of the metric and scalar invariance models were compared with the fit of the configural invariance model. If the factor loadings are not invariant across groups (i.e., metric invariant), then it is possible the scales are not measuring the same constructs in the different groups. If the intercepts of the scale are not invariant across groups (i.e., scalar invariant), then the same score may have different meanings in different groups. As a result, mean comparisons between groups should be interpreted with caution because one group may display higher means as a result of the nonequivalent intercepts of the scale, rather than actual level of prodromal symptoms (F. F. Chen, 2008).

Following convention, if the scale failed to display metric or scalar invariance, we planned to examine the modification indices for clues as to which items in which groups may be responsible for the lack of invariance (e.g., Byrne, Shavelson, & Muthen, 1989). This approach is commonly done in measurement invariance work and is referred to as partial measurement invariance (e.g., Skriner & Chu, 2014; Torres, Miller, & Moore, 2013). Follow-up analyses focused on parameters with modification indices greater than 10.00 (Heene, Hilbert, Freudenthaler, & Buhner, 2012).

For the Total score, we used the DIFFTEST command to calculate a chi-square difference test to determine whether the scalar invariance model fit significantly worse than the configural model. We used the Satorra–Bentler chi-square difference test (Satorra & Bentler, 2001) for the Distress Scale data. We supplemented this analysis with the change in McDonald's noncentrality index (Mc; Mcdonald, 1989) and change in comparative fit index (ΔCFI), as suggested by Meade, Johnson, and Braddy (2008), due to well-known limitations of chi-square-based likelihood ratio tests (e.g., Cheung & Rensvold, 2002). We used cutoffs of 0.02 for Mc and 0.010 for ΔCFI, following the recommendations of Cheung and Rensvold (2002). These cutoffs were originally suggested for MLM. In the current research, we use these suggested cutoffs for both ML and WLS because, to our knowledge, Monte Carlo simulations on appropriate cutoffs for WLS have not been reported.

To determine whether the unequal distribution of sex could have affected the measurement invariance of the results, we also conducted the above mentioned measurement invariance analyses for sex. If the measure has configural, metric, and scalar invariance between sexes, then it is unlikely that the imbalance in sex among ethnicities could have contributed to the invariance results.

Results

Measurement Invariance Analyses

As can be seen in Table 1, the overall single group model fit the data well. In addition, both the configural and the scalar

Table 1. Fit Statistics for Configural, Metric, Scalar, and Scalar Modified Invariance Models by Ethnicity.

Model	χ^2	*df*	RMSEA	90% CI	TLI	CFI	χ^2_{diff} (*df*)	*p*	Mc	ΔCFI
Total Scale										
1. Single Group	1151.17	189	0.041	[0.039, 0.043]	0.953	0.951				
2. Configural	1435.84	756	0.036	[0.033, 0.039]	0.961	0.964				
3. Scalar	11543.76	813	0.036	[0.033, 0.039]	0.961	0.962	133.073 (57)	<.001	0.008	0.002
Distress Scale										
4. Single Group	1928.48	189	0.053	[0.051, 0.055]	0.943	0.936				
5. Configural	1769.93	756	0.044	[0.042, 0.047]	0.883	0.894				
6. Metric	1842.33	816	0.043	[0.040, 0.045]	0.890	0.893	67.909 (60)	0.226	0.002	0.001
7. Scalar	2028.11	876	0.044	[0.041, 0.046]	0.885	0.880	253.52 (120)	<.001	0.020	0.014
8. Modified Scalar	1966.79	872	0.043	[0.040, 0.045]	0.890	0.886	187.019 (116)	<.001	0.012	0.008

Note. df = degrees of freedom; RMSEA = root mean square error of approximation; CI = confidence interval; TLI = Tucker–Lewis index; CFI = comparative fit index; Mc = McDonald's noncentrality index; Model 1 = single group one-factor model; Model 2 = configural model in which the factor loadings and intercepts are free to differ in all groups for the yes/no questions; Model 3 = scalar invariance model in which the factor loadings and intercepts are constrained to be equal across groups for the yes/no questions; Model 4 = single group one-factor distress model; Model 5 = configural model in which the factor loadings and intercepts are free to differ in all groups for the Distress questions; Model 6 = metric invariance model in which the intercepts are free but the factor loadings are constrained to be equal across groups for the Distress questions; Model 7 = scalar invariance model in which the factor loadings and intercepts are constrained to be equal across groups for the Distress questions; Model 8 = modified Scalar invariance model in which the intercepts for the Asian group are freed for Items 8 and 21, and the intercept for both the Asian and White group are freed for Item 17.

Table 2. Fit Statistics for Configural, Metric, Scalar, and Scalar Modified Invariance Models by Sex.

Model	χ^2	*df*	RMSEA	90% CI	TLI	CFI	χ^2_{diff} (*df*)	*p*	Mc	ΔCFI
Total										
1. Configural	1320.947	378	0.040	[0.038, 0.043]	0.952	0.957				
2. Scalar	1373.319	397	0.040	[0.038, 0.043]	0.953	0.956	70.385 (19)	<.001	0.005	0.001
Distress										
3. Configural	1489.453	378	0.044	[0.042, 0.047]	0.888	0.899				
4. Metric	1542.000	398	0.044	[0.042, 0.046]	0.890	0.896	48.354 (20)	<.001	0.005	0.003
5. Scalar	1637.128	418	0.044	[0.042, 0.046]	0.889	0.889	144.981 (40)	<.001	0.015	0.010

Note. df = degrees of freedom; RMSEA = root mean square error of approximation; CI = confidence interval; TLI = Tucker–Lewis index; CFI = comparative fit index; Mc = McDonald's noncentrality index; Model 1 = configural model in which the factor loadings and intercepts are free to differ in all groups for the yes/no questions; Model 2 = scalar invariance model in which the factor loadings and intercepts are constrained to be equal across groups for the yes/no questions; Model 3 = configural model in which the factor loadings and intercepts are free to differ in all groups for the Distress questions; Model 4 = metric invariance model in which the intercepts are free but the factor loadings are constrained to be equal across groups for the Distress questions; Model 5 = scalar invariance model in which the factor loadings and intercepts are constrained to be equal across groups for the Distress questions.

invariance models fit the data well for the PB-Q Total Scale. The scalar invariance model did not fit worse than the configural model according to the ΔCFI and the Mc indices. However, it fit significantly worse according to the chi-square difference test. Given the limitations of the chi-square difference test, and two of the three indices indicating scalar invariance, we interpreted this finding to mean that the Total Scale is scalar invariant. This suggests that the scale is measuring the same construct across groups and that mean score comparisons across groups are meaningful.

As can also be seen in Table 1, the one-factor model for the Distress Scale fit the data well for single group analysis. The configural and metric invariance models also fit the data well. The metric invariance model fit the data just as well as the configural invariance model according to all three fit indices. Thus, we interpreted this to mean that the scale has metric invariance. In contrast, the scalar invariance model fit worse than the configural invariance model according to all three indices. The modification indices suggested that four intercepts should be unconstrained to increase the model fit. These included the intercepts in the Asian group for Item number 8 and Item 21, and the intercepts in both the Asian and White group for Item 17. These three problematic items included Item 8, "Do you feel that other people are watching you or talking about you?" Item 17, "Are your thoughts sometimes so strong that you can almost hear them?" and Item 21, "Do people sometimes find it hard to understand what you are saying?" When these four intercepts were unconstrained, the scalar invariance model fit just as well as the configural invariance model according to the ΔCFI and Mc Indices, but not the chi-square difference test. Thus, we interpreted the Distress Scale as having partial scalar invariance.[1] As can be seen in Table 2, the Total score version of the scale had configural

Table 3. Psychometric Characteristics of the PQ-B Across Ethnic Groups.

	Total scores				*F* or χ^2	Distress scores				*F*
	Asian	White	Hispanic	Multi		Asian	White	Hispanic	Multi	
Total	6.12 (4.69)[a]	5.18 (4.55)[b]	5.25 (4.58)[b]	6.33 (4.67)[a]	9.14*	18.69 (16.42)[a]	14.89 (15.11)[b]	15.58 (15.51)[b,c]	17.71 (15.16)[a,c]	9.75*
PQ1	0.21 (0.41)[a]	0.17 (0.38)[a]	0.19 (0.40)[a]	0.21 (0.40)[a]	3.96	0.75 (1.44)[a]	0.62 (1.37)[a]	0.63 (1.34)[a]	0.74 (1.44)[a]	1.55
PQ2	0.39 (0.49)[a]	0.37 (0.48)[a]	0.40 (0.49)[a]	0.37 (0.48)[a]	1.08	1.19 (1.59)[a]	1.01 (1.44)[a]	1.15 (1.52)[a]	1.02 (1.47)[a]	2.40
PQ3	0.17 (0.38)[a]	0.12 (0.32)[b]	0.16 (0.37)[a,b]	0.14 (0.35)[a,b]	10.27*	0.54 (1.20)[a]	0.36 (1.00)[b]	0.47 (1.13)[a,b]	0.43 (1.10)[a,b]	4.14*
PQ4	0.17 (0.37)[a]	0.17 (0.38)[a]	0.14 (0.34)[a]	0.19 (0.39)[a]	4.68	0.49 (1.13)[a]	0.43 (0.98)[a]	0.39 (1.06)[a]	0.47 (1.03)[a]	0.96
PQ5	0.24 (0.42)[a]	0.19 (0.39)[a,b]	0.16 (0.37)[b]	0.25 (0.43)[a]	15.28*	0.84 (1.54)[a]	0.67 (1.41)[a,b]	0.57 (1.35)[b]	0.93 (1.65)[a]	5.59*
PQ6	0.49 (0.50)[a]	0.38 (0.49)[b]	0.47 (0.50)[a]	0.46 (0.50)[a,b]	21.94*	1.49 (1.66)[a]	1.12 (1.54)[b]	1.31 (1.57)[a,b]	1.34 (1.61)[a,b]	7.83*
PQ7	0.22 (0.41)[a]	0.21 (0.41)[a]	0.18 (0.38)[a]	0.23 (0.42)[a]	3.94	0.54 (1.11)[a]	0.47 (0.98)[a]	0.45 (1.08)[a]	0.52 (1.06)[a]	1.13
PQ8	0.51 (0.50)[a]	0.38 (0.49)[b]	0.38 (0.49)[b,c]	0.47 (0.50)[ac]	44.67*	1.74 (1.82)	1.24 (1.72)	1.28 (1.73)	1.52 (1.76)	NA
PQ9	0.28 (0.45)[a]	0.17 (0.37)[b]	0.26 (0.44)[a]	0.22 (0.42)[a,b]	32.83*	1.00 (1.63)[a]	0.57 (1.29)[b]	0.89 (1.56)[a]	0.79 (1.52)[a,b]	11.43*
PQ10	0.34 (0.48)[a]	0.32 (0.46)[a]	0.29 (0.46)[a]	0.33 (0.47)[a]	4.38	1.07 (1.55)[a]	0.84 (1.37)[b]	0.89 (1.48)[a,b]	0.93 (1.39)[a,b]	4.33
PQ11	0.34 (0.47)[a]	0.34 (0.47)[a]	0.35 (0.48)[a]	0.46 (0.50)[b]	17.20*	1.18 (1.72)[a]	1.08 (1.63)[a]	1.19 (1.73)[a,b]	1.47 (1.77)[b]	3.89
PQ12	0.37 (0.48)[a]	0.28 (0.45)[b]	0.26 (0.44)[b]	0.39 (0.49)[a]	28.70*	1.36 (1.83)[a]	1.10 (1.78)[b]	0.95 (1.66)[b]	1.41 (1.86)[a]	8.09*
PQ13	0.17 (0.37)[a]	0.14 (0.35)[a]	0.16 (0.37)[a]	0.18 (0.39)[a]	2.05	0.59 (1.33)[a]	0.50 (1.25)[a]	0.48 (1.17)[a]	0.57 (1.28)[a]	1.18
PQ14	0.42 (0.49)[a]	0.36 (0.48)[b,c]	0.35 (0.48)[c]	0.44 (0.49)[a,b]	13.33*	1.35 (1.66)[a]	1.09 (1.55)[b]	1.08 (1.57)[b]	1.34 (1.63)[a,b]	5.67*
PQ15	0.28 (0.45)[a]	0.28 (0.45)[a]	0.24 (0.43)[a]	0.37 (0.48)[b]	15.93*	0.73 (1.27)[a,b]	0.66 (1.18)[a,b]	0.60 (1.18)[a]	0.85 (1.30)[b]	2.87
PQ16	0.19 (0.39)[a,b]	0.18 (0.39)[a,b]	0.14 (0.35)[b]	0.23 (0.43)[a]	11.59*	0.60 (1.30)[a]	0.57 (1.27)[a]	0.44 (1.17)[a]	0.67 (1.31)[a]	2.50
PQ17	0.22 (0.42)[a]	0.29 (0.45)[b,c]	0.22 (0.41)[a,c]	0.31 (0.47)[b]	19.76*	0.65 (1.29)	0.79 (1.34)	0.60 (1.22)	0.83 (1.32)	NA
PQ18	0.51 (0.50)[a]	0.42 (0.49)[b]	0.43 (0.49)[b]	0.59 (0.49)[c]	34.90*	1.71 (1.78)[a]	1.32 (1.68)[b]	1.40 (1.73)[b]	1.89 (1.76)[a]	12.01*
PQ19	0.13 (0.34)[a]	0.11 (0.31)[a]	0.11 (0.31)[a]	0.13 (0.33)[a]	3.36	0.43 (1.11)[a]	0.33 (0.94)[a]	0.35 (1.06)[a]	0.36 (1.01)[a]	1.68
PQ20	0.13 (0.34)[a]	0.10 (0.31)[a]	0.09 (0.29)[a]	0.13 (0.34)[a]	7.19	0.43 (1.14)[a]	0.28 (0.87)[b]	0.28 (0.96)[a,b]	0.40 (1.07)[a,b]	4.21
PQ21	0.41 (0.49)[a]	0.27 (0.44)[b]	0.29 (0.45)[b,c]	0.36 (0.48)[a,c]	48.36*	1.33 (1.68)	0.83 (1.43)	0.90 (1.53)	1.10 (1.62)	NA

Note. PQ-B = Prodromal Questionnaire–Brief. Means with the same superscript do not significantly differ from each other.
*$p < .05$.

and scalar invariance between sexes, and the Distress score had configural, metric, and scalar invariance between sexes, both according to the Mc and ΔCFI but not the chi-square difference test. This suggests that the PQ-B scores represent the same construct and latent level of attenuated psychotic symptoms between sexes. Moreover, it suggests that the different rates of sex among groups likely did not affect the measurement invariance results.

Overall, the measurement invariance results suggest that the Total Scale is measuring the same construct in these different groups, and that the manifest scores represent that same level of the latent construct among groups. Moreover, the finding that only 3 of the 21 items lack scalar invariance suggests that Distress means can be compared across groups. However, these comparisons should be interpreted with caution.

Mean Comparisons

To examine whether the Total Scale means differed among groups, we compared the Total Scale means with a one-way analysis of variance (ANOVA). To determine whether the proportion of yes/no answers differed across groups, we conducted a series of chi-squares for each item comparing across all four groups and among each pair of groups. A Bonferroni correction was applied to adjust the alpha level for multiple comparisons.

As can be seen in Table 3, the Asian and Multiracial groups had higher total mean scores than the Hispanic and White groups, and group differences emerged for 13 of the 21 total items. Differences on the items followed the same pattern as the Total scores, with the White and Hispanic groups tending to have lower scores than the Asian and Multiracial groups.[2] The White and Hispanic groups differed from each other on only two items (6 and 9). On both items, Hispanic participants were more likely to answer affirmatively. Likewise, Asian and Multiracial participants differed on only 4 of the 21 items, (11, 15, 17, and 18). Asian participants had higher scores than White participants on eight items (3, 6, 8, 9, 12, 14, 18, and 21) but lower scores on one item (17). Asian participants had higher scores than Hispanic participants on six items (5, 8, 12, 14, 18, and 21). White participants had lower scores than Multiracial participants on six items (8, 11, 12, 15, 18, and 21). Finally, Hispanic participants had lower scores than Multiracial participants on eight items (5, 11, 12, 14, 15, 16, 17, and 18).

For Distress scores, we ran a series of ANOVAs for the Total score and each individual item, again applying the Bonferroni correction for multiple comparisons. As described above, the Distress Scale did not display scalar invariance, which means that mean comparisons should be interpreted with caution. We did not compute ANOVA scores or pair comparisons for the three items that were

responsible for the lack of scalar invariance (i.e., Items 8, 17, and 21). However, we report the comparisons for the total Distress score because only three of the 21 items (i.e., 14.2%) were found to be problematic, which is below the recommended threshold of 20% for partial scalar invariance (Byrne et al., 1989). Nevertheless, this comparison should be interpreted with caution. Overall, the Asian group reported more distress than the White and Hispanic groups, while the Multiracial group reported more distress than the White group. This pattern held true for the individual items as well, with the Asian and Multiracial groups tending to have higher scores than the White and Hispanic groups. The Asian group had lower scores than the Multiracial group on two items (11 and 15), higher scores than the White group on seven items (3, 9, 10, 12, 14, 18, and 20), and higher scores than the Hispanic group on four items (5, 12, 14, and 18). The White group had lower scores on than the Hispanic group on one item (9), and lower scores than the Multiracial group on three items (11, 12, and 18). Finally the Hispanic group had lower scores than the Multiracial group on four items (5, 12, 15, and 18).

Discussion

The results of the current research suggest that the PQ-B Total score has scalar invariance among Asian, White, Hispanic, and Multiracial participants. Since scalar invariance involves constraining both the intercepts and factor loadings to be equal across groups, this implies that the Total score also has metric invariance across groups, which only requires the factor loadings to be equal across groups. Similarly, the Distress Scale displayed metric invariance across these groups and partial scalar invariance. The lack of full scalar invariance can be attributed to three problematic items, which is below the recommended cutoff of 20% of items (Byrne et al., 1989). This is significant for the use of the PQ-B for at least two reasons. First, the finding of scalar invariance for the Total score and metric invariance for the Distress Scale suggests that the scale is measuring the same construct in these different groups. This is important because it confirms that previous research using the PQ-B in Asian, White, Hispanic, and Multiracial undergraduates yields similarly interpretable scores across groups. As mentioned, many studies have used the PQ-B in undergraduates, and although the majority were White participants in most studies, there were Asian, Hispanic, and Multiracial participants in these studies. These results suggest that their scores and correlations with other variables can be interpreted in the same way as the White participants.

The second finding related to measurement invariance was that the Total Scale had scalar invariance and the Distress Scale had partial scalar invariance. This finding suggests that the scores represent the same level of latent attenuated psychotic symptoms in these four groups. Although there were differences in PQ-B Total and Distress scores among groups, these results suggest that they represent actual differences in symptoms rather than measurement factors such as differences in demand characteristics, response styles, or item interpretations between groups. Since the measurement invariance findings indicated that the scale measures the same construct and the scores represent the same level of latent attenuated psychotic symptoms across groups, this suggests that there are not cultural differences that pertain to the diagnostic relevance of the symptoms. Previous researchers have suggested creating different cut points for different ethnic groups based on differences in mean levels (Chmielewski et al., 1995). Based on these results, we would not recommend creating different cut scores for Asian, White, Hispanic, and Multiracial participants with the PQ-B.

A second major finding of the current research was that there were differences among the ethnic groups in Total and Distress scores. Asian and Multiracial participants had higher PQ-B scores than the White and Hispanic participants. This finding is consistent with research showing elevated rates of attenuated psychotic-symptoms in ethnic minorities in general (Sharpley & Peters, 1999), and Asian American and Multiracial participants in particular (Cicero, 2015; Schiffman, 2004). In contrast, this finding is in inconsistent with previous research that has found Hispanic participants tend to have higher psychotic-like experience scores than White participants (Chmielewski et al., 1995). In previous work comparing Hispanic and White participants, Chmielewski et al. (1995) used the Wisconsin Schizotypy Scales. However, the measurement invariance of the Wisconsin Schizotypy Scales has not been established between White and Hispanic participants, and many of the items have been shown to have differential item functioning between White participants and other ethnic minorities (Winterstein et al., 2011). Thus, the previous findings may be related to the psychometric properties of that scale, rather than true differences in psychotic-like experiences. The findings with respect to measurement invariance suggest that these differences are related to actual differences in the experience of attenuated psychotic symptoms, rather than the psychometric properties of the PQ-B.

The current research identified three items as having different intercepts between groups. Item 8, "Do you feel that other people are watching you or talking about you?" may tap into the way that one views oneself in society (e.g., independent or interdependent self-construal; Markus & Kitayama, 1991). Interdependent societies, like many Eastern cultures, afford greater group cohesion and social support, but may lead to increased feelings of surveillance and cautiousness toward others (Kitayama, Duffy, & Uchida, 2007). In contrast to a cultural explanation, Item 21, "Do people sometimes find it hard to understand what you are saying?"

may reflect acculturative difficulties among first generation Asian Americans, such as English language acquisition.

In addition to these differences, some of the differences in item intercepts may be related to well-documented, group-level differences in response style (Iwata, 2014). For example, some response styles, such as acquiescence, may account for the varying intercept in Item 17, "Are your thoughts sometimes so strong that you can almost hear them?" The Distress questions may activate cultural scripts related to self-enhancement (i.e., minimizing one's negative or neutral characteristics) or self-effacement (i.e., emphasizing one's negative or neutral characteristics; Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997) when responding about oneself.

In the current research, we operationalized ethnicity as self-reported ethnic identity in which participants reported their ethnicities on a single question. Although the current results provide evidence for the measurement invariance of the PQ-B Total scores and for 18 of the 21 Distress scores, it does not provide insight into why those three items were not invariant and why there were mean differences in global scores and many individual items. As mentioned, researchers have suggested several potential cultural mechanisms that confer risk for attenuated psychotic symptoms, such as poor strength of ethnic identity, racial discrimination, acculturation, and social adversity among others. None of these measures were included in the current research. Future research could explore these relations in more detail by including measures of these factors to see if they are correlated with PQ-B scores or original items.

In addition to not including these additional measures, one limitation of the current work is that no measures were included to assess for careless or invalid responding. It is possible that some participants did not answer the questions thoughtfully or truthfully. Moreover, no measures of past psychiatric history were included, which could have helped provide context for the current results. Additionally, the current study is the first to our knowledge to administer the scale to American undergraduates online. Previous paper-and-pencil tests in similar samples have found slightly lower means ($M = 4.70$ for the Total score, and $M = 13.19$ for the Distress score; DeVylder, Thompson, et al., 2015). It is possible that the results would have been different if the study were administered with a paper-and-pencil format. The current research did not include a large enough sample of African American and Native American participants to test measurement invariance in these groups. Future research could examine whether the PQ-B is also invariant in African Americans and Native Americans.

One clear limitation of the current research is the use of an undergraduate sample. This means that the results may not generalize to noncollege young adults from the general population, or to students who do not take psychology courses while in college. College students tend to have higher socioeconomic status and education than the general population and are functioning well enough to be enrolled in college at the time of participation. In addition to differences in baseline psychotic symptoms, one could argue that college students may be less likely to experience some of the negative aspects of ethnic minority status that are associated with psychotic-like experiences, such as social adversity and discrimination. However, a longline of research suggests that ethnic minority college students experience discrimination, racism, micro-aggressions, acculturative stress, imposter feelings, minority status stress, and increased social adversity at similar levels to noncollege attending peers (e.g., Blume, Lovato, Thyken, & Denny, 2012; Cokley, 2002; Juang et al., 2016; Metzger, Cooper, Ritchwood, Onyeuku, & Griffin, 2016; Okazaki, 1997; Steele & Aronson, 1995).

At the same time, undergraduates may have more in common with the general population than has been previously thought. It is unclear if empirical data support the conclusion that people with college students are psychologically healthier than noncollege attending peers. The National Epidemiologic Survey of Alcohol and Related Conditions found that there were no differences between college students and nonstudents in the same age range in total psychopathology (Blanco et al., 2008). Other research has found that college students experience attenuated psychotic symptoms at relatively high rates (Cicero et al., 2014; Loewy et al., 2007). According to the National Center for Education Statistics (2016), there are over 20 million currently enrolled college students in the United States, representing 40% of 18 to 24-year olds. More than two thirds of high school graduates enroll in college the following year (U.S. Bureau of Labor Statistics, 2016). Thus, even if the current results are only generalizable to people who attend college, the results would still be meaningful to a large percentage of the population, and may be useful to clinicians and researchers working with undergraduate populations. Nevertheless, one area for future research is to examine whether there is measurement invariance in the PQ-B in a random sample drawn from diverse geographic areas in the United States and potentially globally. Since the sample was drawn from U.S. colleges, the current results may not generalize to the same ethnic groups in samples drawn from international populations. In addition to noncollege attending young adults, the results of the current research may not generalize to clinical samples of outpatients in either psychology clinics or prodromal high-risk clinics. Future research could examine the measurement invariance of this scale in White, Asian, Multiracial, and Hispanic participants drawn from both settings.

Another limitation of the current research is that the potential of any screening measure for psychosis risk in the general population may be limited. Previous research has found that the majority of high scorers on self-report measures do not go

on to develop psychotic-spectrum disorders, even after long periods of follow-up (Chapman, Chapman, Kwapil, Eckblad, & Zinser, 1994; Kwapil, Miller, Zinser, Chapman, & Chapman, 1997). One proposed strategy for these screening measures is to screen a large portion of the population and bring people above a predetermined cut score into a clinic for more intensive assessment with a structured interview (F. Chen et al., 2014), such as the Structured Interview for Prodromal Syndromes (Miller et al., 2003). It is not clear how effective this strategy will prove to be because (a) many people who exceed the cut score have been shown to be "false positives" in a follow-up interview (Kline & Schiffman, 2014) and (b) the majority of patients who meet criteria for attenuated psychosis syndrome do not go on to develop psychosis (Cannon et al., 2008). Thus, there are two opportunities in this strategy to people to be incorrectly determined to be at risk for psychosis. Future research could include studies in which general population samples are screened for psychosis risk and then followed longitudinally to determine if the PQ-B is a useful tool as a screening device in such populations. This research could determine appropriate cut scores for these populations to decrease false positives and increase specificity. Although the scales had measurement invariance among groups, it is still possible that the predictive validity of PQ-B scores could vary among the groups.

In summary, the current research provides further support for the reliability and validity of PQ-B scores in White, Asian, Multiracial, and Hispanic undergraduates. The Total scores had configural and scalar invariance across these four groups and the Distress scores had configural, metric, and partial scalar invariance in these groups. These results suggest that the scale measures the same constructs in these groups and that scale scores represent the same latent level of attenuated psychotic symptoms among groups. Critically, this major finding suggests that the scale can be used in these diverse groups and that previous research using the scale in these groups is valid.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. In addition to running these analyses with MLM estimation, we collapsed Categories 4 and 5 in all groups and ran the model with WLSMV estimation. All three models, configural, metric, and scalar had good model fit (RMSEAs < .05, and CFI and TLI > .95). Both the metric and scalar models fit worse than the configural model according to the DIFFTEST chi-square difference test, but not worse according to the Mc index (both < .02) and ΔCFI (both < .01). Thus, with WLSMV, the PQ-B has metric and scalar invariance within these groups. We report and interpret the MLM analyses in the main body of the text because it suggests a more conservative interpretation and there may be valuable information lost by grouping 4 and 5 together because some ethnic groups may be less likely to select extreme values, which would indicate a lack of measurement invariance.
2. To examine whether these differences could be accounted for by different proportions in sex, we ran an analysis of covariance that partialled out variance associated with sex. The pattern of results for the Total and Distress scores was identical in these analyses.

References

Addington, J., & Heinssen, R. (2012). Prediction and prevention of psychosis in youth at clinical high risk. *Annual Review of Clinical Psychology*, *8*, 269-289. doi:10.1146/annurev-clinpsy-032511-143146

Addington, J., Van Mastrigt, S., Hutchinson, J., & Addington, D. (2002). Pathways to care: Help seeking behaviour in first episode psychosis. *Acta Psychiatrica Scandinavica*, *106*, 358-364.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Anglin, D. M., Lighty, Q., Greenspoon, M., & Ellman, L. M. (2014). Racial discrimination is associated with distressing subthreshold positive psychotic symptoms among US urban ethnic minority young adults. *Social Psychiatry and Psychiatric Epidemiology*, *49*, 1545-1555. doi:10.1007/s00127-014-0870-8

Benton, S. A., Robertson, J. M., Tseng, W.-C., Newton, F. B., & Benton, S. L. (2003). Changes in counseling center client problems across 13 years. *Professional Psychology: Research and Practice*, *34*, 66-72. doi:10.1037/0735-7028.34.1.66

Blanco, C., Okuda, M., Wright, C., Hasin, D. S., Grant, B. F., Liu, S. M., & Olfson, M. (2008). Mental health of college students and their non-college-attending peers: Results from the National Epidemiologic Study on Alcohol and Related Conditions. *Archives of General Psychiatry*, *65*, 1429-1437. doi:10.1001/archpsyc.65.12.1429

Blume, A. W., Lovato, L. V., Thyken, B. N., & Denny, N. (2012). The relationship of microaggressions with alcohol use and anxiety among ethnic minority college students in a historically White institution. *Cultural Diversity & Ethnic Minority Psychology*, *18*, 45-54. doi:10.1037/a0025457

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Burnett-Zeigler, I., Bohnert, K. M., & Ilgen, M. A. (2013). Ethnic identity, acculturation and the prevalence of lifetime psychiatric disorders among Black, Hispanic, and Asian adults in the U.S. *Journal of Psychiatric Research*, *47*, 56-63. doi:10.1016/j.jpsychires.2012.08.029

Byrne, B. M., Shavelson, R. J., & Muthen, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456-466. doi:10.1037/0033-2909.105.3.456

Cannon, T. D., Cadenhead, K., Cornblatt, B., Woods, S. W., Addington, J., Walker, E., . . . Heinssen, R. (2008). Prediction of psychosis in youth at high clinical risk: A multisite longitudinal study in North America. *Archives of General Psychiatry*, *65*, 28-37. doi:10.1001/archgenpsychiatry.2007.3

Chakraborty, A., & McKenzie, K. (2002). Does racial discrimination cause mental illness? *British Journal of Psychiatry*, *180*, 475-477.

Chapman, L. J., Chapman, J. P., Kwapil, T. R., Eckblad, M., & Zinser, M. C. (1994). Putatively psychosis-prone subjects 10 years later. *Journal of Abnormal Psychology*, *103*, 171-183. doi:10.1037/0021-843x.103.2.171

Chen, F., Wang, L., Heeramun-Aubeeluck, A., Wang, J., Shi, J., Yuan, J., & Zhao, X. (2014). Identification and characterization of college students with Attenuated Psychosis Syndrome in China. *Psychiatry Research*, *216*, 346-350. doi:10.1016/j.psychres.2014.01.051

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, *95*, 1005-1018. doi:10.1037/a0013193

Chen, W. J., Hsiao, C. K., & Lin, C. C. H. (1997). Schizotypy in community samples: The three-factor structure and correlation with sustained attention. *Journal of Abnormal Psychology*, *106*, 649-654. doi:10.1037/0021-843X.106.4.649

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

Chmielewski, P. M., Fernandes, L. O., Yee, C. M., & Miller, G. A. (1995). Ethnicity and gender in scales of psychosis proneness and mood disorders. *Journal of Abnormal Psychology*, *104*, 464-470. doi:10.1037/0021-843x.104.3.464

Cicero, D. C. (2015). Measurement invariance of the Schizotypal Personality Questionnaire in Asian, Pacific Islander, White, and multiethnic populations. *Psychological Assessment*, *28*, 351-361. doi:10.1037/pas0000180

Cicero, D. C., Martin, E. A., Becker, T. M., Docherty, A., & Kerns, J. G. (2014). Correspondence between psychometric and high clinical risk for psychosis in an undergraduate population. *Psychological Assessment*, *26*, 901-915. doi:10.1037/a0036432

Cokley, K. O. (2002). Ethnicity, gender, and academic self-concept: A preliminary examination of academic disidentification and implications for psychologists. *Cultural Diversity & Ethnic Minority Psychology*, *8*, 378-388.

Cooper, S., Klugman, J., Heimberg, R. G., Anglin, D. M., & Ellman, L. M. (2016). Attenuated positive psychotic symptoms and social anxiety: Along a psychotic continuum or different constructs? *Psychiatry Research*, *235*, 139-147. doi:10.1016/j.psychres.2015.11.027

Denenny, D., Thompson, E., Pitts, S. C., Dixon, L. B., & Schiffman, J. (2015). Subthreshold psychotic symptom distress, self-stigma, and peer social support among college students with mental health concerns. *Psychiatric Rehabilitation Journal*, *38*, 164-170. doi:10.1037/prj0000124

DeVylder, J. E., Jahn, D. R., Doherty, T., Wilson, C. S., Wilcox, H. C., Schiffman, J., & Hilimire, M. R. (2015). Social and psychological contributions to the co-occurrence of sub-threshold psychotic experiences and suicidal behavior. *Social Psychiatry and Psychiatric Epidemiology*, *50*, 1819-1830. doi:10.1007/s00127-015-1139-6

Devylder, J. E., Oh, H. Y., Yang, L. H., Cabassa, L. J., Chen, F. P., & Lukens, E. P. (2013). Acculturative stress and psychotic-like experiences among Asian and Latino immigrants to the United States. *Schizophrenia Research*, *150*, 223-228. doi:10.1016/j.schres.2013.07.040

DeVylder, J. E., Thompson, E., Reeves, G., & Schiffman, J. (2015). Psychotic experiences as indicators of suicidal ideation in a non-clinical college sample. *Psychiatry Research*, *226*, 489-493. doi:10.1016/j.psychres.2015.02.007

Fusar-Poli, P., Borgwardt, S., Bechdolf, A., Addington, J., Riecher-Rössler, A., Schultze-Lutter, F., . . . Yung, A. (2013). The psychosis high-risk state: A comprehensive state-of-the-art review. *JAMA Psychiatry*, *70*, 107-120. doi:10.1001/jamapsychiatry.2013.269

Gibson, L. E., Anglin, D. M., Klugman, J. T., Reeves, L. E., Fineberg, A. M., Maxwell, S. D., . . . Ellman, L. M. (2014). Stress sensitivity mediates the relationship between traumatic life events and attenuated positive psychotic symptoms differentially by gender in a college population sample. *Journal of Psychiatric Research*, *53*, 111-118. doi:10.1016/j.jpsychires.2014.02.020

Häfner, H., an der Heiden, W., Behrens, S., Gattaz, W. F., Hambrecht, M., Löffler, W., . . . Stein, A. (1998). Causes and consequences of the gender difference in age at onset of schizophrenia. *Schizophrenia Bulletin*, *24*, 99-113.

Heene, M., Hilbert, S., Freudenthaler, H. H., & Buhner, M. (2012). Sensitivity of SEM fit indexes with respect to violations of uncorrelated errors. *Structural Equation Modeling*, *19*, 36-50. doi:10.1080/10705511.2012.634710

Hunt, J., & Eisenberg, D. (2010). Mental health problems and help-seeking behavior among college students. *Journal of Adolescent Health*, *46*, 3-10. doi:10.1016/j.jadohealth.2009.08.008

Iwata, N. (2014). Cultural distinctiveness in response bias. In M. F. Dollard, A. Shimazu, R. Bin Nordin, P. Brough & M. R. Tuckey (Eds.), *Psychosocial factors at work in the Asia Pacific* (pp. 55-88). New York, NY: Springer.

Juang, L. P., Moffitt, U., Kim, S. Y., Lee, R. M., Soto, J. A., Hurley, E., . . . Whitborne, S. K. (2016). Cognitive reappraisal and expressive suppression: Links to racial-ethnic discrimination and adjustment among Latino/a and Asian-heritage college students. *Journal of Adolescence*, *53*, 21-33. doi:10.1016/j.adolescence.2016.08.012

Karlsen, S., Nazroo, J. Y., McKenzie, K., Bhui, K., & Weich, S. (2005). Racism, psychosis and common mental disorder among ethnic minority groups in England. *Psychological Medicine*, *35*, 1795-1803. doi:10.1017/s0033291705005830

Kitayama, S., Duffy, S., & Uchida, Y. (2007). Self as cultural mode of being. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 136-174). New York, NY: Guilford Press.

Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology*, *72*, 1245-1267. doi:10.1037/0022-3514.72.6.1245

Kline, E., & Schiffman, J. (2014). Psychosis risk screening: A systematic review. *Schizophrenia Research*, *158*, 11-18. doi:10.1016/j.schres.2014.06.036

Kline, E., Wilson, C., Ereshefsky, S., Tsuji, T., Schiffman, J., Pitts, S., & Reeves, G. (2012). Convergent and discriminant validity of attenuated psychosis screening tools. *Schizophrenia Research*, *134*, 49-53. doi:10.1016/j.schres.2011.10.001

Kwapil, T. R., Miller, M. B., Zinser, M. C., Chapman, J., & Chapman, L. J. (1997). Magical ideation and social anhedonia as predictors of psychosis proneness: A partial replication. *Journal of Abnormal Psychology*, *106*, 491-495.

Loewy, R. L., Bearden, C. E., Johnson, J. K., Raine, A., & Cannon, T. D. (2005). The Prodromal Questionnaire (PQ): Preliminary validation of a self-report screening measure for prodromal and psychotic syndromes. *Schizophrenia Research*, *79*, 117-125.

Loewy, R. L., Johnson, J. K., & Cannon, T. D. (2007). Self-report of attenuated psychotic experiences in a college population. *Schizophrenia Research*, *93*, 144-151. doi:10.1016/j.schres.2007.02.010

Loewy, R. L., Pearson, R., Vinogradov, S., Bearden, C. E., & Cannon, T. D. (2011). Psychosis risk screening with the Prodromal Questionnaire–Brief Version (PQ-B). *Schizophrenia Research*, *129*, 42-46. doi:10.1016/j.schres.2011.03.029

Markus, H. R., & Kitayama, S. (1991). Culture and self: Implications for cognition, emotion, and motivation. *Psychological Review*, *92*, 224-253.

Marshall, M., Lewis, S., Lockwood, A., Drake, R., Jones, P., & Croudace, T. (2005). Association between duration of untreated psychosis and outcome in cohorts of first-episode patients: A systematic review. *Archives of General Psychiatry*, *62*, 975-983. doi:10.1001/archpsyc.62.9.975

McDonald, R. P. (1989). An index of goodness-of-fit based on noncentrality. *Journal of Classification*, *6*, 97-103.

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. doi:10.1037/0021-9010.93.3.568

Metzger, I. W., Cooper, S. M., Ritchwood, T. D., Onyeuku, C., & Griffin, C. B. (2016). Profiles of African American college students' alcohol use and sexual behaviors: Associations with stress, racial discrimination, and social support. *Journal of Sex Research*, 1-12. doi:10.1080/00224499.2016.1179709

Miller, T. J., McGlashan, T. H., Rosen, J. L., Cadenhead, K., Cannon, T., Ventura, J., . . . Woods, S. W. (2003). Prodromal assessment with the Structured Interview for Prodromal Syndromes and the Scale of Prodromal Symptoms: Predictive validity, interrater reliability, and training to reliability. *Schizophrenia Bulletin*, *29*, 703-715.

Mittal, V. A., Dean, D. J., Pelletier, A., & Caligiuri, M. (2011). Associations between spontaneous movement abnormalities and psychotic-like experiences in the general population. *Schizophrenia Research*, *132*, 194-196. doi:10.1016/j.schres.2011.06.028

Moller, P., & Husby, R. (2000). The initial prodrome in schizophrenia: Searching for naturalistic core dimensions of experience and behavior. *Schizophrenia Bulletin*, *26*, 217-232.

Muthen, L. K., & Muthen, B. O. (1998-2016). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

National Center for Education Statistics. (2016). *Back to school statistics*. Retrieved from http://nces.ed.gov/fastfacts/display.asp?id=372

Okazaki, S. (1997). Sources of ethnic differences between Asian American and White American college students on measures of depression and social anxiety. *Journal of Abnormal Psychology*, *106*, 52-60.

Satorra, A., & Bentler, P. A. (2001). A scale difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507-514.

Schiffman, J. (2004). Schizotypal traits in a non-clinical sample from Hawai'i. *Pacific Health Dialog*, *11*, 84-86.

Sharpley, M. S., & Peters, E. R. (1999). Ethnicity, class and schizotypy. *Social Psychiatry and Psychiatric Epidemiology*, *34*, 507-512. doi:10.1007/s001270050168

Skriner, L. C., & Chu, B. C. (2014). Cross-ethnic measurement invariance of the SCARED and CES-D in a youth sample. *Psychological Assessment*, *26*, 332-337. doi:10.1037/a0035092

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797-811. doi:10.1037/0022-3514.69.5.797

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. doi:10.1037/a0030436

U.S. Bureau of Labor Statistics. (2016). *College enrollment and work activity of 2015 high school graduates*. Retrieved from http://www.bls.gov/news.release/hsgec.nr0.htm

van Os, J., Hanssen, M., Bijl, R. V., & Vollebergh, W. (2001). Prevalence of psychotic disorder and community level of psychotic symptoms: An urban-rural comparison. *Archives of General Psychiatry*, *58*, 663-668.

Wicks, S., Hjern, A., Gunnell, D., Lewis, G., & Dalman, C. (2005). Social adversity in childhood and the risk of developing psychosis: A national cohort study. *American Journal of Psychiatry*, *162*, 1652-1657. doi:10.1176/appi.ajp.162.9.1652

Winterstein, B. P., Ackerman, T. A., Silvia, P. J., & Kwapil, T. R. (2011). Psychometric properties of the Wisconsin Schizotypy Scales in an undergraduate sample: Classical test theory, item response theory, and differential item functioning. *Journal of Psychopathology and Behavioral Assessment*, *33*, 480-490. doi:10.1007/s10862-011-9242-9

Yung, A. R., Nelson, B., Baker, K., Buckby, J. A., Baksheev, G., & Cosgrave, E. M. (2009). Psychotic-like experiences in a community sample of adolescents: Implications for the continuum model of psychosis and prediction of schizophrenia. *Australian & New Zealand Journal of Psychiatry*, *43*, 118-128. doi:10.1080/00048670802607188

Yung, A. R., Yuen, H. P., McGorry, P. D., Phillips, L. J., Kelly, D., Dell'Olio, M., . . . Buckby, J. (2005). Mapping the onset of psychosis: The Comprehensive Assessment of At-Risk Mental States. *Australian & New Zealand Journal of Psychiatry*, *39*, 964-971.