Title

Measurement Invariance of the Persistent Complex Bereavement Disorder Checklist With Respect to Youth Gender, Race, Ethnicity, and Age

Journal

Journal of Traumatic Stress, 33(5)

ISSN

0894-9867

Authors

Hill, Ryan M
Dodd, Cody
Oosterhoff, Benjamin
et al.

Publication Date

2020-10-01

DOI

10.1002/jts.22560

Peer reviewed

Journal of Traumatic Stress
October 2020, *33,* 850–856

International Society for Traumatic Stress Studies

Measurement Invariance of the Persistent Complex Bereavement Disorder Checklist With Respect to Youth Gender, Race, Ethnicity, and Age

Ryan M. Hill,[1] Cody Dodd,[1] Benjamin Oosterhoff,[2] Christopher M. Layne,[3,4] Robert S. Pynoos,[3,4] Mary Beth Staine,[5] and Julie B. Kaplow[1]

[1]Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA
[2]Department of Psychology, Montana State University, Bozeman, Montana, USA
[3]UCLA/Duke University National Center for Child Traumatic Stress, Los Angeles, California, USA
[4]Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, USA
[5]Bo's Place, Houston, Texas, USA

The Persistent Complex Bereavement Disorder (PCBD) Checklist was constructed to facilitate the developmentally sensitive assessment of proposed PCBD criteria in bereaved children and adolescents 8–18 years of age. Initial analyses of the PCBD Checklist provided support for the hypothesized two-factor model. The purpose of the present study was to evaluate the measurement invariance of the PCBD Checklist with respect to gender (boys and girls), race/ethnicity (White, Black, and Hispanic youth), and age (school age, preadolescent, and adolescent youth). Participants were 594 youth (50.4% female) aged 7–18 years ($M = 11.91$, $SD = 2.80$) who were evaluated as part of standard care at a community-based grief support center. Youth self-identified as Hispanic ($n = 184$, 30.8%), non-Hispanic white ($n = 179$, 30.0%), and African American/Black ($n = 136$, 22.8%). A series of stepwise, multigroup confirmatory factor analyses provided evidence in support of the PCBD Checklist's measurement invariance for all three groups concerning configural invariance, metric invariance, and scalar invariance. These results suggest that PCBD Checklist Criterion B and C scores are measuring similar latent variables, to a similar degree, across gender, race/ethnicity, and age. Establishing the cross-group equivalence of the PCBD Checklist is an important endorsement of its generalizability and clinical utility in that it can be administered to diverse populations with confidence that it is measuring proposed PCBC diagnostic criteria similarly across subgroups.

The inclusion of persistent complex bereavement disorder (PCBD) as a provisional (i.e., candidate) disorder in the appendix of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-5*; American Psychiatric Association [APA], 2013) served as a call to action for rigorous examination of the proposed PCBD criteria. The construction and validation of developmentally sensitive assessment tools is an essential early step in ensuring that proposed diagnostic symptoms both accurately reflect the phenomenological experiences of bereaved youth and facilitate the rigorous empirical evaluation of PCBD as a diagnostic entity (Kaplow et al., 2012; Nader & Layne, 2009). The PCBD Checklist (Kaplow et al. 2018; Layne et al., 2014), which assesses PCBD symptoms among youth ages 8–18 years, was designed to address this need. An initial psychometric evaluation of the PCBD Checklist in a sample of bereaved youth supported its reliability, validity, and two-factor structure (Kaplow et al., 2018). However, further examination of the measure's psychometric properties is needed to evaluate the generalizability of the measure across different populations and settings. Of particular interest is the empirical examination of the structure of the PCBD Checklist across diverse bereaved samples. Accordingly, the primary aim of the present study was to empirically examine the measurement invariance of the PCBD Checklist with respect to gender, race and ethnicity, and age. In addition, potential latent group differences regarding mean PCBD Checklist scores were examined.

In its appendix, the *DSM-5* established five criteria for verifying a proposed diagnosis of PCBD (APA, 2013). Criterion A requires that the individual has experienced the death of someone with whom they had a close relationship. Criteria B and C comprise the primary symptom clusters, which were generated to include features that have been consistently associated with maladaptive grieving in prior research and rationally

The project was supported by funding from the New York Life Foundation (PI: Julie B. Kaplow).

Correspondence concerning this article should be addressed to Ryan M. Hill, Department of Pediatrics, Baylor College of Medicine, Houston, TX. E-mail: Ryan.Hill@bcm.edu

DOI: 10.1002/jts.22560

partitioned into two primary symptom clusters. Criterion B symptoms include (a) separation distress, defined as persistent and intense yearning, longing, sorrow, and preoccupation with the deceased, and (b) preoccupation with the circumstances of the death. Criterion C symptoms include (c) reactive distress in response to the death, including difficulty accepting the death, difficulty reminiscing, and excessive avoidance of loss reminders, and (d) disruptions in personal and social identity, including feeling like part of oneself has died. Prior research regarding the PCBD Checklist has provided evidence for the content validity and two-factor structure of Criterion B and C (Kaplow et al., 2018). Criterion D establishes that the identified symptoms result in clinically significant distress or impairment, and Criterion E requires that symptoms be out of proportion to or inconsistent with cultural, religious, or age-appropriate norms.

The PCBD diagnosis also includes a "traumatic bereavement specifier," which is endorsed when the death occurred via homicide or suicide. The extant literature suggests that a small but substantial subset of bereaved youth (i.e., approximately 10%) report developing a syndrome distinct from normal grief reactions that corresponds with some of the proposed PCBD criteria (Dillen et al., 2009; Layne et al., 2001; Melhem et al., 2008, 2011). The findings from a more recent study demonstrated higher rates of PCBD (i.e., approximately 18%) among a diverse sample of youth (Kaplow et al., 2018).

The PCBD Checklist (Kaplow et al., 2018; Layne et al., 2014) was constructed to facilitate the developmentally sensitive assessment of proposed PCBD criteria in bereaved children and adolescents between 8 and 18 years of age. The measure includes a one-page, eight-item cover sheet containing items regarding the child's relationship to the deceased (i.e., establishing Criterion A), the amount of time elapsed since the death, and the manner of death (i.e., establishing whether the traumatic bereavement specifier is indicated). The cover sheet is followed by a 39-item self-report survey that assesses the primary symptom domains as well as functional impairment. The PCBD Checklist was designed using best-practice test construction procedures (DeVellis, 2012). Specifically, the development of the PCBD Checklist involved (a) careful examination of developmental manifestations of proposed PCBD criteria (Kaplow et al., 2012), (b) generation of a developmentally informed item pool, (c) evaluation of the item pool by an expert panel, (d) focus-group evaluation to determine the comprehensibility and acceptability of the items, and (e) evaluation of the psychometric properties of the resulting item pool, including discriminant groups, convergent validity, and discriminant validity (Kaplow et al., 2018). Confirmatory factor analyses of the measure provided support for the hypothesized two-factor model, as compared with a single factor model, in the form of latent constructs corresponding to PCBD Criterion B and C symptom clusters (Kaplow et al., 2018; Layne et al., 2014). In addition, bereaved youth who met the PCBD diagnostic criteria according to PCBD Checklist scoring guidelines had elevated scores on established measures of posttraumatic stress and depressive symptoms relative to youth who did not meet the criteria (Kaplow et al., 2018).

To further evaluate the psychometric properties of the PCBD Checklist, we utilized measurement invariance testing to evaluate whether the constructs assessed are measured similarly across demographic groups. Measurement invariance describes three sets of multiple-group confirmatory factor analyses that are carried out in a stepwise fashion (Brown, 2006). In the first step, a test of configural invariance was used to examine whether the number of factors and patterns of loadings on the factors were statistically similar across groups. Configural invariance is assumed if the configural invariance model provides an acceptable fit to the data in all groups. As applied to the PCBD Checklist, evidence of configural invariance would imply that the underlying factor structure of the assessment tool, divided into Criterion B and Criterion C symptoms clusters, is adequate for all groups and that the same items are associated with each symptom cluster across groups.

In the second step, we used a test of metric invariance to compared groups regarding the strength of the associations between indicator items and their corresponding factors (i.e., the factor loadings; Brown, 2006). When supported, metric invariance indicates that the strength of the association of each item to the factor is the same across groups, suggesting that the definition of the factor is similar across groups. Metric invariance is established by setting equal factor loadings across groups. If the model fit is substantially worsened by the addition of this constraint (Kline, 2011) as compared to the configural invariance model, then metric invariance is not supported. Metric invariance would imply that the reflective indicators (i.e., PCBD symptoms) contribute similarly to the Criteria B and C latent factors across groups.

In the third step, a test of scalar invariance was used to examine group differences in item thresholds or intercepts for categorical or continuous variables, respectively (Brown, 2006). Scalar invariance indicates that scores on a latent construct across groups reflect similar levels of the latent phenomenon. Scalar invariance is established by holding item thresholds or intercepts equal across groups. If this constraint significantly worsens model fit as compared to the metric invariance model, scalar invariance cannot be assumed. Taken together, configural, metric, and scalar invariance constitute strong evidence that members of different subgroups grieve in similar ways as measured by the PCBD Checklist.

The purpose of the present study was to build upon past research that has examined the psychometric properties of the PCBD Checklist to facilitate the continued rigorous evaluation of the proposed PCBD criteria among bereaved youth. Measurement invariance of the PCBD Checklist was evaluated with respect to gender (boys and girls), race/ethnicity (White, Black, and Hispanic youth), and age (school age, preadolescent, and adolescent youth). Finally, group differences in mean PCBD factor scores were evaluated with respect to gender, race/ethnicity, and age.

Table 1
Clinical Characteristics of Sample Groups

Variable	Sample size	Age (years)		PCBD Criterion B score[a]		PCBD Criterion C score[b]	
		M	*SD*	*M*	*SD*	*M*	*SD*
Full sample	594	11.93	2.79	24.41	6.25	60.59	19.49
Gender							
Boys	292	11.84	2.63	23.67	6.43	58.40	19.36
Girls	299	12.03	2.93	25.15	6.00	62.74	19.47
Race/ethnicity							
White	179	11.96	2.89	23.01	6.34	56.24	18.61
Black	135	11.67	2.57_a	25.54	5.90	63.42	19.51
Hispanic	183	12.52	2.85_a	25.27	5.91	62.83	19.05
Age group							
School age	221	9.06	0.81_c	24.41	5.90	63.32	19.55
Preadolescent	191	11.90	0.82_c	24.20	6.46	58.84	19.95
Adolescent	182	15.45	1.20_c	24.62	6.48	59.11	18.64

Note. Shared subscripts indicate significant differences at $p < .05$.
[a]525 youth (88.4%) met suggested cutoffs for PCBD Criterion B (at least one of four symptoms present). [b]308 youth (51.9%) met suggested cutoffs for PCBD Criterion C (at least six of 12 symptoms present).

Method

Participants and Procedure

The present study was approved by the institutional review board of Baylor College of Medicine. Participants were 594 youth between 8 and 18 years of age (M = 11.93, SD = 2.79) who were evaluated as part of standard care at a community-based grief support center that provides grief support, resources, and referrals, as needed. Youth and families seeking care were asked to participate as part of a standard intake evaluation for services; youth assent and parent or guardian consent were obtained before evaluation. A total of 612 families were approached for participation; 18 declined and 594 (97.1%) provided consent for participation. Upon completion of the evaluation, youth were directed internally to grief support groups or referred to a bereavement-focused outpatient clinic, as appropriate. Youth were remunerated $10 (USD) for their participation.

Youth (50.4% female) self-identified their race/ethnicity as Hispanic (n = 183, 30.8%), non-Hispanic White (n = 179, 30.0%), African American/Black (n = 135, 22.8%), Asian American/Pacific Islander (n = 18, 3.0%), Native American (n = 5, 0.8%), and other, biracial, or multiracial (n = 75, 12.6%). Two individuals did not report their gender. The relationships between the youth and the deceased included the participant's father or stepfather (n = 268, 44.9%), mother or stepmother (n = 144, 24.2%), brother or sister (n = 120, 20.1%), grandparent or great-grand parent (n = 46, 7.7%), aunt or uncle (n = 6, 1.0%), and other (e.g., nephew, cousin, close friend; n = 14, 2.1%). The most common cause of death was persisting illness (e.g., cancer, lung disease; n = 229, 38.6%), followed by an accident (n = 115, 19.4%), an illness that made them die suddenly (e.g., a heart attack; n = 111, 18.7%), suicide (n = 48, 8.1%), murder or homicide (n = 43, 7.3%), drug overdose (n = 11, 1.8%), other (n = 11, 1.8%), and cause of death unknown to the respondent (n = 25, 4.2%). More than one third (n = 231, 38.8%) had experienced multiple deaths (M = 1.59, SD = 1.59, range: 1–7). The average time since the death was 8.76 months (SD = 16.05). The overall sample was divided into subgroups by gender (boys, n = 292, girls, n = 299), race/ethnicity (n = 179 White, n = 135 Black, and n = 183 Hispanic), and age group (school age [8–10 years], n = 221; preadolescent [11–13 years], n = 191; and adolescent [14–18 years], n = 182). Clinical characteristics of each sample group are provided in Table 1.

Measures

Persistent complex bereavement disorder

The PCBD Checklist is a 39-item measure designed to assess *DSM-5* provisional PCBD criteria in bereaved children and adolescents (8–18 years of age) and identify youth at an elevated risk for clinically significant maladaptive grief reactions (Layne et al., 2014). Items assess past-month symptoms and are rated on a 5-point Likert-type scale ranging from 0 (*not at all*) to 4 (*all the time*). Example items include "I feel so sad about losing ____ that my heart aches" (Criterion B, separation distress) and "I feel like when ____ died, a big part of me died too" (Criterion C, disruptions in personal and social identity). The PCBD Checklist has exhibited good convergent, discriminant, and discriminant-group validity as well as developmental appropriateness and clinical utility in a diverse sample of bereaved children and adolescents (Kaplow et al., 2018). In the present sample, 525 youth (88.4%) met the suggested cutoff for PCBD Criterion B (at least one of four symptoms present) and 308 youth (51.9%) met the suggested cutoff for PCBD Criterion C (at least six of 12 symptoms present).

Data Analysis

All factor analyses and omega calculations (Muthén & Muthén, 1998–2017) were conducted using M*plus* (Version 8.2). Observed mean values were calculated using SPSS Statistics (Version 25). A small percentage of data (0.08%) were missing for 13 items, with no case missing more than two items. To account for missing data, a "pairwise present" approach was used (Asparouhov & Muthen, 2010). Given the ordinal nature of the item response options, a mean- and variance-adjusted weighted least squares estimator (WLSMV) was employed for all analyses, with item indicators specified as categorical (Bowen & Masa, 2015). Model fit was evaluated based on traditionally accepted standards (Barret, 2007) and utilized the root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker–Lewis index (TLI). Acceptable model fit required CFI and TLI values .90 or higher and an RMSEA value of .08 or lower (Little, 2013). Due to the large size of the total sample and subgroups and limitations of the chi-square difference test, significant chi-square difference tests were not considered unequivocally supportive of less restrictive models (Meade et al., 2008; Milfont & Fischer, 2010). Instead, CFI differences were used to evaluate model comparisons, with a CFI difference of .010 or higher taken as evidence of significant worsening of model fit (Cheung & Rensvold, 2002). In other words, change in CFI was used because it is less sensitive to large samples than the chi-square difference test.

Per Brown's (2006) recommendation, tests of measurement invariance were carried out in the following sequence: First, confirmatory factor analyses (CFA) were conducted to evaluate the factor structure and model fit separately for each group. Multiple group CFAs were then conducted to evaluate configural invariance (i.e., identical factor structure), metric invariance (i.e., the equivalence of factor loadings), and scalar invariance (i.e., the equivalence of indicator thresholds). Strict invariance (i.e., the equivalence of factor and error variances) went untested as it is commonly untenable in social science research (Bowen & Masa, 2015). Measurement invariance models were evaluated using the CONFIGURAL METRIC SCALAR command in M*plus*, which tests all three levels of invariance within a single analysis. Omega values for internal consistency reliability estimates were calculated for both the overall sample and individual groups. Latent group means were examined by fixing the mean of a reference group to zero within the scalar invariance models and evaluating differences between group means. For comparison across three groups, this process was repeated with an alternative group identified as the reference group, to allow for a complete set of cross-group comparisons.

Results

Factor Reliability Estimates

Internal consistency reliability for PCBD Criterion B symptoms was good for the overall sample, $\omega = .82$, with moderate variations across groups, $\omega\text{s} = .76–.85$. For PCBD Criterion C symptoms, reliability was good for the overall sample, $\omega = 0.92$, with minor variations across groups, $\omega\text{s} = .91–.94$. The internal consistencies for each factor and each subgroup are provided in the Online Supplementary Materials.

Measurement Invariance

Prior to evaluating measurement invariance, independent CFAs were performed to ensure that the model adequately fit each subsample. Standardized factor loadings for the full sample are provided in Table 3; factor loadings for each subgroup are provided in the Online Supplementary Materials. Model fit statistics are provided in Table 2. All models showed an acceptable fit. Model fit was similar for boys and girls; White, Black, and Hispanic youth; and school-age, preadolescent, and adolescent youth.

Fit statistics and the results of difference tests are shown in Table 2. Configural invariance models demonstrated acceptable CFI values, ranging from .943 to .946, and TLI values, ranging from .939 to .944; and low RMSEA values, ranging from .060 to .063. This indicated that configural invariance models fit the data adequately across invariance categories. Concerning metric invariance models, fit indices showed acceptable fit with this constraint as indicated by a change in CFI value less than .010 (range: $< .001–.002$). Of note, RMSEA values and TLI values stayed the same or improved across most of the increasingly restrictive models. Thus, individual item factor loadings demonstrated invariance. Regarding scalar invariance, fit indices showed acceptable fit with this constraint, as indicated by a change in CFI of less than .010, as compared to the metric invariance models, for which the range was $< .001$ to .009. Thus, the item thresholds demonstrated invariance, and the two-factor solution of the PCBD Checklist met the criteria for scalar invariance with respect to gender, race/ethnicity, and age groups.

Mean Differences in PCBD Checklist Scores

Based on the establishment of full scalar invariance across groups, comparisons of latent group means were evaluated. These analyses examine group mean differences, accounting for sources of measurement error at the item level. The latent variable was centered so tests of differences in mean values could be conducted. Girls reported significantly higher PCBD scores than boys regarding items related to both Criterion B, $M_{\text{diff}} = 0.17$, $SE = 0.06$, $p = .005$; and Criterion C, $M_{\text{diff}} = 0.14$, $SE = 0.05$, $p = .003$. Black youth reported significantly higher PCBD scores than White youth regarding items related to both Criterion B, $M_{\text{diff}} = 0.35$, $SE = 0.09$, $p < .001$; and Criterion C, $M_{\text{diff}} = 0.23$, $SE = 0.07$, $p = .001$. Hispanic youth

Table 2
Fit Statistics of the Single-Group Confirmatory Factor Analysis and Configural, Metric, and Scalar Invariance Models

Variable	χ^2	N	df	p	RMSEA	90% CI	CFI	TLI	$\Delta\chi^2(df)$	p	ΔCFI
					Single-group CFA models						
Full sample	1,371.18	594	376	< .001	.067	[.063, .070]	.937	.932	–	–	–
Boys	838.64	292	376	< .001	.065	[.059, .071]	.940	.935	–	–	–
Girls	792.66	299	376	< .001	.061	[.055, .067]	.947	.943	–	–	–
White	624.42	179	376	< .001	.061	[.052, .069]	.944	.940	–	–	–
Black	546.69	135	376	< .001	.058	[.047, .068]	.941	.936	–	–	–
Hispanic	644.41	183	376	< .001	.062	[.054, .071]	.944	.940	–	–	–
School age	620.57	221	376	< .001	.054	[.047, .062]	.945	.941	–	–	–
Preadolescent	686.68	191	376	< .001	.066	[.058, .074]	.953	.949	–	–	–
Adolescent	682.52	182	376	< .001	.067	[.059, .075]	.943	.938	–	–	–
					Measurement invariance, by gender						
Configural invariance	1,631.59	591	752	< .001	.063	[.059, .067]	.944	.939	–	–	–
Metric invariance	1,647.91	591	779	< .001	.061	[.057, .066]	.944	.942	37.66 (27)	.084	< .001
Scalar invariance	1,689.68	591	864	< .001	.057	[.053, .061]	.947	.950	118.84 (85)	.009	.003
					Measurement invariance, by race						
Configural invariance	1,801.12	497	1,128	< .001	.060	[.055, .065]	.946	.941	–	–	–
Metric invariance	1,849.06	497	1,182	< .001	.058	[.053, .063]	.946	.944	72.07 (54)	.051	< .001
Scalar invariance	2,021.34	497	1,352	< .001	.055	[.050, .060]	.946	.951	244.26 (170)	< .001	< .001
					Measurement invariance, by age						
Configural invariance	1,987.38	594	1,128	< .001	.062	[.058, .066]	.948	.944	–	–	–
Metric invariance	2,070.95	594	1,182	< .001	.062	[.057, .066]	.946	.945	113.53 (54)	< .001	.002
Scalar invariance	2,369.76	594	1,352	< .001	.062	[.058, .066]	.939	.945	385.69 (170)	< .001	.009

Note. Chi-square difference tests were computed using the M*plus* shortcut for difference testing. RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker–Lewis index; CFA = confirmatory factor analysis.

reported significantly higher PCBD scores than White youth for Criterion B, $M_{\text{diff}} = 0.29$, $SE = 0.09$, $p = .001$; and Criterion C: $M_{\text{diff}} = 0.21$, $SE = 0.07$, $p = .001$. However, the mean scores for Black and Hispanic youth did not significantly differ from each other: Criterion B, $M_{\text{diff}} = -0.06$, $SE = 0.09$, $p = .525$; Criterion C, $M_{\text{diff}} = -0.02$, $SE = 0.07$, $p = .801$.

Preadolescent youth reported significantly lower scores than school-age youth for Criterion C, $M_{\text{diff}} = -0.12$, $SE = 0.05$, $p = .031$; but not Criterion B, $M_{\text{diff}} = -0.07$, $SE = 0.07$, $p = .311$. Adolescent youth reported significantly lower scores than school-age youth for Criterion C, $M_{\text{diff}} = -0.12$, $SE = 0.06$, $p = .026$, but not Criterion B, $M_{\text{diff}} = -0.03$, $SE = 0.07$, $p = .640$. Mean scores for preadolescent and adolescent youth did not significantly differ from each other: Criterion B, $M_{\text{diff}} = 0.04$, $SE = 0.07$, $p = .604$; Criterion C, $M_{\text{diff}} = -0.01$, $SE = 0.06$, $p = .916$.

Discussion

This was the first to examine the cross-group psychometric properties of the PCBD Checklist. Specifically, we examined the measurement invariance of the PCBD Criterion B and C symptom scales with regard to gender (boys and girls), race/ethnicity (White, Black, and Hispanic youth), and age (school age, preadolescent, and adolescent) in a sample of bereaved youth assessed at a community-based bereavement support center. The findings from a set of stepwise, multigroup confirmatory factor analyses provided evidence in support of the PCBD Checklist's measurement invariance for all three groups. First, equivalence was demonstrated in terms of configural invariance, which indicated that the same group of items load onto a two-factor structure of PCBD Criteria B and C symptoms for all groups. Next, evidence of metric invariance indicated comparable magnitudes of factor loadings across groups. Finally, evidence of scalar invariance indicated that observed item thresholds were equivalent across groups.

The analyses revealed group differences in the PCBD Checklist latent mean scores with regard to gender, race/ethnicity, and age subgroups. Specifically, for both Criterion B and Criterion C, girls reported higher scores than boys, and Black and Hispanic youth reported higher scores than White youth. Adolescent and preadolescent youth reported lower Criterion C scores than school-age youth, but these groups did not significantly differ on Criterion B scores. Of note, these differences corresponded to score differences of less than 0.35 points on the 5-point scale. These results suggest that, after accounting for measurement error at the item level, there were measurable group differences in PCBD symptom severity. Additional research is needed to examine potential explanatory factors that may account for these differences across gender, race/ethnicity, and

Table 3
Standardized and Unstandardized Factor Loadings for the Full Sample

Item	Unstandardized coefficient	*SE*	Standardized coefficient	*SE*
Criterion B				
X1	1.000	0.000	.674	.025
X2	0.772	0.055	.520	.034
X6	0.943	0.049	.636	.028
X9	0.931	0.051	.627	.029
X13	1.079	0.056	.727	.028
X19	1.199	0.052	.808	.021
X24	1.086	0.053	.731	.024
Criterion C				
X3	1.000	0.000	.555	.031
X4	1.214	0.075	.674	.025
X7	1.322	0.077	.733	.022
X8	1.137	0.074	.631	.028
X10	1.245	0.074	.691	.025
X12	1.287	0.077	.714	.024
X14	1.315	0.078	.730	.023
X16	1.178	0.074	.653	.026
X18	1.242	0.077	.689	.025
X20	1.342	0.080	.745	.021
X21	1.337	0.079	.742	.023
X23	1.148	0.077	.637	.027
X26	0.702	0.082	.390	.039
X27	1.365	0.082	.758	.020
X28	1.096	0.071	.608	.028
X29	0.907	0.077	.503	.035
X31	0.963	0.077	.534	.034
X32	1.367	0.081	.759	.020
X33	1.261	0.078	.700	.025
X34	0.916	0.074	.508	.033
X35	1.220	0.079	.555	.025
X36	1.245	0.079	.674	.026

age groups. For example, heightened exposure to violent deaths among Black and Hispanic youth living in low–socioeconomic status neighborhoods might account for symptom differences. In addition, younger children, particularly parentally bereaved school-age youth, may be more at risk for certain maladaptive grief reactions given their increased dependence on caregivers at this age (Kaplow et al., 2012).

Together, these results suggest that the PCBD Checklist Criterion B and C scores are measuring similar latent variables, to a similar degree, across subgroups based on gender, race/ethnicity, and age. Establishing the cross-group equivalence of the PCBD Checklist is an important endorsement of its generalizability and clinical utility in that it can be administered in diverse populations with confidence that it is measuring proposed PCBD diagnostic criteria similarly across subgroups. Scalar measurement invariance across demographic groups suggests that the PCBD construct measured using the PCBD Checklist is statistically equivalent across groups; group means can be compared, and the associations between PCBD scores and other constructs can be examined for all groups at the same time.

The results of this study should be considered in the context of its limitations. Although the sample was large and diverse in terms of the three racial/ethnic groups compared, limited enrollment precluded analyses of measurement invariance for other racial or ethnic groups (e.g., American Indian/Alaskan Native, Asian, Native Hawaiian/Other Pacific Islander, and multiracial). Future research is needed to determine whether the current results extend to these and other cultural subgroups. Similarly, although measurement invariance was demonstrated for English-speaking Hispanic participants, future research is needed to examine the measurement invariance across the English- and Spanish-language versions of the PCBD Checklist. Along these lines, the current research was conducted with a support-seeking sample of bereaved children and adolescents, and further research is needed to examine the psychometric properties of the PCBD Checklist with a naturalistic sample of bereaved youth.

There is some debate as to the nuances of model evaluation and model comparison for the examination of measurement invariance. Chi-squared difference tests have been criticized as being highly sensitive to sample size, producing an overly restrictive threshold for evaluating differences in model fit, whereas CFI differences may be too lenient a threshold (Meade et al., 2008). Meade and colleagues (2008) suggest a more conservative cutoff of less than -.002, although Little (2013) suggests such a cutoff may be overly conservative (for a review of measurement invariance conventions and their limitations, see Putnick & Bornstein, 2016). In the present analyses, we relied on CFI differences for interpreting outcomes and used the less-conservative .01 threshold (Cheung & Rensvold, 2002). Further, although our model fit indices are adequate according to Little's (2013) threshold, there is some debate about whether acceptable model fit should require a minimum CFI of .95 (Hu & Bentler, 1999).

Although measurement invariance is a critical component of evaluating test bias (Reynolds & Suzuki, 2012), further research should examine potential differential associations between the Criterion B and C scores and their respective associations with external measures, including indicators of functional impairment in various life domains, such as school, family, and peer relations, as well as risky behavior and other indicators of mental distress. Future research should also expand on the present findings by evaluating whether the discriminant and convergent validity of the PCBD Checklist varies across demographic subgroups. Another promising area of research that would expand upon the current results involves examining the association between Criterion B and C scores and other contextual variables assessed within the PCBD Checklist, including the cause of death, relationship to the deceased, time elapsed since the death, or whether the PCBD traumatic bereavement specifier is endorsed. The performance of the PCBD Checklist should also be evaluated in the context of single versus multiple deaths, including both multisequential deaths and multiple deaths occurring as part of a single evaluation (e.g., a disaster or plane crash), as well as violent versus nonviolent deaths (e.g., those for which the traumatic bereavement specifier is indicated). These results would both extend the clinical utility of the PCBD Checklist and inform risk prediction models aimed at identifying youth at risk for maladaptive grief reactions.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Open Practices Statement

Due to the ongoing nature of the research project, the data utilized in this study are not currently available for public use; the study was not preregistered.

References

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

Asparouhov, T., & Muthén, B. (2010). *Weighted least squares estimation with missing data*. https://www.statmodel.com/download/GstrucMissingRevision.pdf

Barrett, P. (2007). Structural equation modelling: adjudging model fit. *Personality and Individual Differences*, *42*, 815–824. https://doi.org/10.1016/j.paid.2006.09.018

Bowen, N. K., & Masa, R. D. (2015). Conducting measurement invariance tests with ordinal data: A guide for social work researchers. *Journal of the Society for Social Work and Research*, *6*, 229–249. https://doi.org/10.1086/681607

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

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

DeVellis, R. F. (2012). *Scale development: Theory and applications*. Sage Publications.

Dillen, L., Fontaine, J. R. J., & Verhofstadt-Denève, L. (2009). Confirming the distinctiveness of complicated grief from depression and anxiety among adolescents. *Death Studies*, *33*, 437–461. https://doi.org/10.1080/07481180902805673

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

Kaplow, J. B., Layne, C. M., Oosterhoff, B., Goldenthal, H., Howell, K. H., Wamser-Nanney, R., Burnside, A., Calhoun, K., Marbury, D., Johnson-Hughes, L., Kriesel, M., Staine, M. B., Mankin, M., Portorter-Howard, L., & Pynoos, R. (2018). Validation of the Persistent Complex Bereavement Disorder (PCBD) Checklist: A developmentally informed assessment tool for bereaved youth. *Journal of Traumatic Stress*, *31*, 244–254. https://doi.org/10.1002/jts.22277

Kaplow, J. B., Layne, C. M., Pynoos, R. S., Cohen, J., & Lieberman, A. (2012). *DSM-5* diagnostic criteria for bereavement-related disorders in children and adolescents: Developmental considerations. *Psychiatry*, *75*(3), 242–265. https://doi.org/10.1521/psyc.2012.75.3.243

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Publications.

Layne, C. M., Kaplow, J. B., & Pynoos, R. S. (2014). *Persistent Complex Bereavement Disorder (PCBD) Checklist—Youth Version 1.0.* University of California, Los Angeles.

Layne, C. M., Pynoos, R. S., Saltzman, W. R., Arslangaic, B., Black, M., Savjak, N., Popović, N., Duraković, E., Mušić, M., Ćampara, N., Djapo, N., & Houston, R. (2001). Trauma/grief-focused group psychotherapy: School-based postwar intervention with traumatized Bosnian adolescents. *Group Dynamics: Theory, Research, and Practice*, *5*, 277–290. https://doi.org/10.1037/1089-2699.5.4.277

Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford Press.

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*(3), 568–592. https://doi.org/10.1037/0021-9010.93.3.568

Melhem, N. M., Porta, G., Shamseddeen, W., Walker Payne, M., & Brent, D. A. (2011). Grief in children and adolescents bereaved by sudden parental death. *Archives of General Psychiatry*, *68*(9), 911–919. https://doi.org/10.1001/archgenpsychiatry.2011.101

Melhem, N. M., Walker, M., Moritz, G., & Brent, D. A. (2008). Antecedents and sequelae of sudden parental death in offspring and surviving caregivers. *Archives of Pediatrics & Adolescent Medicine*, *162*(5), 403–410. https://doi.org/10.1001/archpedi.162.5.403

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, *3*(3), 111–130. https://doi.org/10.21500/20112084.857

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nader, K. O., & Layne, C. M. (2009). Maladaptive grieving in children and adolescents: Discovering developmentally-linked differences in the manifestation of grief. *Traumatic Stress Points*, *23*(5), 12–16. https://doi.org/10.1037/e518302011-004

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Reynolds, C. R., & Suzuki, L. A. (2012). Bias in psychological assessment. In I. B. Weiner, J. R. Graham, & J. A. Naglieri (Eds.), *Handbook of psychology* (2nd ed; pp. 82–113). https://doi.org/10.1002/9781118133880.hop210004