ELSEVIER

Measurement invariance of caregiver support following sexual abuse across age, relationship, and English-Spanish language☆

Caitlin Rancher [a],[*], Daniel W. Smith [a], Rosaura Orengo-Aguayo [a], Mindy Jackson [b], Ernest N. Jouriles [c]

[a] National Crime Victims Research Treatment Center, Medical University of South Carolina, Charleston, SC, USA
[b] Dallas Children's Advocacy Center, Dallas, TX, USA
[c] Department of Psychology, Southern Methodist University, Dallas, TX, USA

ABSTRACT

Background: The Maternal Self-report Support Questionnaire (MSSQ) is among the most rigorously evaluated measures of caregiver support following child sexual abuse, but there is a paucity of data on the factor structure and variance of the MSSQ across diverse groups of caregivers and their children.
Objective: The present study examined the factor structure and measurement invariance of the MSSQ across 386 non-offending caregivers following a disclosure of child sexual abuse.
Participants and setting: Data were collected from non-offending caregivers at two Child Advocacy Centers ($n = 277$; $n = 109$) in the United States.
Methods: Caregivers completed the MSSQ and assessments of child age, caregiver-child relationship, and caregiver preferred language.
Results: Confirmatory factor analyses replicated the original two-factor structure, with the emotional support and blame/doubt subscales emerging as distinct factors. Multigroup confirmatory factor analyses showed measurement invariance across child age and caregiver-child relationship (mother vs. another caregiver). Evidence of partial invariance was found for caregiver preferred language. Comparisons of scores indicated caregiver support varied by child age and caregiver preferred language.
Conclusions: Overall, findings suggest the MSSQ can be used to measure caregiver support across caregivers with children of different ages and both mothers and non-mothers, but caution should be practiced in interpreting mean-level differences between English- and Spanish-speaking caregivers.

Support from a nonoffending caregiver is theorized as an important predictor of children's recovery from sexual abuse (Finkelhor & Browne, 1985; Spaccarelli, 1994). In line with theory, measures of caregiver support have been found to relate to child adjustment in dozens of studies (Bolen & Gergely, 2015; Elliott & Carnes, 2001), and caregiver support is considered a key component of many empirically supported treatments following child sexual abuse (e.g., Trauma-Focused Cognitive Behavior Therapy; Cohen et al., 2017;

Parent-Child Interaction Therapy; Thomas & Zimmer-Gembeck, 2012). However, there is a dearth of rigorous research on the measurement of caregiver support following child sexual abuse, and a paucity of evidence of the utility of measures of caregiver support across different samples of nonoffending caregivers. Examining whether caregiver support can be operationalized and assessed similarly across different samples of caregivers is a critical step for being able to draw meaningful conclusions about any observed differences across samples of caregivers; for example, such research helps to establish whether differences in scores across groups of caregivers reflect meaningful differences between groups, rather than differences in caregivers' interpretation or understanding of a measure.

In much of the research on caregiver support following child sexual abuse, the measurement of caregiver support suffers from both conceptual and methodological weaknesses (see Bolen & Gergely, 2015; Wamser-Nanney, 2017, for reviews). For example, many researchers have measured caregiver support with single-items assessing caregivers' "*belief that the abuse occurred*" (Wamser-Nanney, 2017); this is a prohibitively narrow conceptualization and assessment of the construct of support, and it is inconsistent with theory that conceptualizes caregiver support following child sexual abuse as a multi-dimensional construct (e.g., Spaccarelli, 1994). Multi-dimensional measures of caregiver support do exist, but most have either unknown or limited data regarding their psychometric properties, and they have not been evaluated for use across different samples of caregivers. For example, the clinician-administered Parental Response to Abuse Disclosure Scale (PRADS; Everson et al., 1989) conceptualizes caregiver support as a multi-dimensional construct across four domains: emotional support, belief, action against the perpetrator, and use of professional services. Each domain is assessed by a single item coded by a trained examiner. Unfortunately, no additional psychometric data has been established for the PRADS. The Parental Support Questionnaire (PSQ; Mannarino & Cohen, 1996) assesses emotional support and blame. Although there is evidence of internal consistency and test-retest reliability in prior research with the PSQ (Cohen & Mannarino, 2000), no additional psychometric data are available.

The Maternal Self-report Support Questionnaire (MSSQ; Smith et al., 2010) is one of the few theoretically derived measures of caregiver support following child sexual abuse that has undergone rigorous psychometric evaluation. Consistent with theory (Spaccarelli, 1994), the MSSQ was developed to assess caregiver support across two domains: emotional support and blame/doubt. The initial evaluation of the MSSQ included exploratory factor analyses that supported the two-factor model of emotional support and blame/doubt (Smith et al., 2010). There was also evidence of internal consistency, convergent validity with multi-reporter ratings of maternal support, and concurrent criterion validity with children's internalizing and externalizing adjustment problems (Smith et al., 2010). Since the initial evaluation of the MSSQ, this measure has been used in a number of studies on families of children who have been sexually abused, including several that have documented associations between maternal support and child adjustment (e.g., see Wamser-Nanney, 2017 for review).

However, despite these strengths, the MSSQ has only been rigorously evaluated among samples of primarily English-speaking mothers. Additionally, there has been minimal research evaluating whether the MSSQ measures caregiver support the same way across caregivers with children of different ages, different caregiver-child relationships, or caregivers' preferred language. The absence of rigorous measurement research using different samples of caregivers and children hampers the ability of researchers and clinicians to confidently distinguish between true, meaningful group-level differences and measurement confounds. The present study therefore seeks to examine the factor structure and measurement invariance of the MSSQ across theoretically salient groups of child age (school aged [6–11 years] vs. adolescent [12–17 years]), caregiver-child relationship (mothers vs. non-mothers), and caregiver preferred language (English vs. Spanish).

1.1. Caregiver support and child age

The shift between school aged childhood, typically considered between 6 and 11 years of age, and adolescence, between 12 and 17 years, is a developmental period marked by transitions in social relationships, as well as cognitive ability and self-regulation (Colman et al., 2006). School aged children typically rely on the guidance and support from their caregivers (Colman et al., 2006). In contrast, adolescents often seek greater autonomy from caregivers and form close, supportive relationships with individuals outside of the home (Jackson & Goossens, 2020). Previous research has found differences across child age in caregiver emotional support and blame/doubt following child sexual abuse disclosure. Specifically, when a child is 11 years old or younger, mothers are more likely to believe their child's report of sexual abuse and provide emotional support to their child following a disclosure, compared to when a child is older (Bolen & Lamb, 2002; Sufredini et al., 2020). Similarly, in the original evaluation of the MSSQ, mothers endorsed higher levels of emotional support for younger children, as compared to older children (Smith et al., 2010); however, child age was not correlated with maternal blame/doubt. Given these previous findings, it is important to examine whether the construct of caregiver support differs across groups of caregivers with school aged children (ages 6 to 11 years) and adolescents (ages 12 to 17 years). Items that best measure caregiver emotional support and blame/doubt for younger children may not work as well for older children and vice versa, challenging our interpretation of mean level differences in the absence of research on measurement invariance.

1.2. Caregiver support and caregiver-child relationship

Caregiver support is most typically operationalized and measured as maternal support, and many researchers have found that following child sexual abuse, mothers' provision of warm, emotional support is associated with child well-being (Elliott & Carnes, 2001). However, many children receive care and support from non-mother caregivers, such as fathers, grandparents, aunts, and other relatives (U.S. Department of Commerce, Census Bureau, 2020). There are important social differences between mothers and non-mother caregivers. For example, mothers generally tend to be more sensitive to the instrumental and emotional needs of their

children, whereas male caregivers typically are more involved in play, achievement, and recreation (see Lewis & Lamb, 2003; Rothenberg et al., 2020, for reviews). Nevertheless, several systematic reviews have asserted that support from non-mother caregivers, including fathers (Rollè et al., 2019), grandparents (Dunifon, 2013), and other kinship care placements (Winokur et al., 2018) play a critical role in promoting child adjustment.

Despite the prevalence and importance of non-mothers in children's lives, empirical examinations of non-mother caregivers following child sexual abuse are sparse. A recent meta-analysis reviewed the handful of exceptions and found that the child's relationship to the caregiver, operationalized as mothers vs. non-mothers, moderated the association between caregiver support and child adjustment (Bolen & Gergely, 2015). Specifically, the association between caregiver support and child adjustment was stronger in studies that included non-mother caregivers, compared to studies that only included mothers. This finding suggests non-mother caregivers may play a critical role in improving our understanding of the association between caregiver support and child adjustment. However, it might be necessary to operationalize caregiver support differently for mothers and non-mothers.

1.3. Caregiver support and caregiver preferred language

There are numerous challenges and benefits to establishing measurement invariance in caregiver support across English- and Spanish-speaking caregivers. Hispanics and Latinos comprise the largest minority group in the United States, and over 13% of the population speaks Spanish at home (U.S. Department of Commerce, Census Bureau, 2019). Past research has evidenced that parenting behaviors can vary greatly by race, ethnicity, and culture (see Sangawi et al., 2015, for review). Further, previous examinations of the MSSQ have found that following a sexual abuse disclosure, Black and Hispanic mothers endorsed higher levels of blame/doubt compared to White mothers (Smith et al., 2010; Wamser-Nanney, 2017). However, most of this research relies on conceptualizations and measures of parenting based on English-speaking, traditional Americanized values. From a construct validity perspective, measurement noninvariance across languages can occur for many different reasons, including: items are linguistically translated, but differ in concepts or meaning; participants may interpret items across different cultural norms; and, certain groups may avoid or use more extreme response choices on Likert-scale measures (Chen & West, 2008). To our knowledge there are no empirical comparisons of caregiver support following child sexual abuse across caregivers who speak English and Spanish. Given the importance of valid mean comparisons, and the prevalence of Spanish-speaking caregivers, establishing that the measurement of caregiver support is invariant across caregiver's preferred language is critical.

1.4. Present study

The present study extends the measurement literature on caregiver support following child sexual abuse by examining the factor structure, measurement invariance, and mean-level differences of the MSSQ across a diverse group of caregivers and their children. First, we conducted confirmatory factor analyses on the items from the MSSQ and hypothesized that the two-factor model assessing emotional support and blame/doubt would replicate in a larger, heterogenous sample of caregivers. Second, we sought to evaluate whether this best-fitting factor structure replicated across diverse groups of caregivers and their children. Specifically, we examined the measurement invariance of the MSSQ across groups of child age (school aged [6–11 years] vs. adolescent [12–17 years]), caregiver-child relationship (mothers vs. non-mothers), and caregiver preferred language (English vs. Spanish). Despite the developmental, social, and cultural differences across these groups, we hypothesized that the measurement of caregiver support, and the multigroup confirmatory factor analyses, would indicate the MSSQ factor model replicated across all groups. After evaluating invariance in the measurement of caregiver support, we explored whether mean and item levels of caregiver emotional support and blame/doubt varied across child age, caregiver-child relationship, and caregiver preferred language.

2. Method

2.1. Participants and procedures

Data for the present study were collected from two samples of non-offending caregivers following a disclosure of child sexual abuse and presentation at two children's advocacy centers (CAC). All study procedures were approved by university Institutional Review Boards and the Review Board of the Sample 2 CAC. All caregivers were informed their decision to participate in the research study would not impact the services they received at the CAC.

2.1.1. Sample 1

Participants provided data as part of a routine assessment conducted by family advocates at a CAC in the southern United States. These assessments typically occurred within a couple of weeks following a forensic evaluation at the same CAC. Inclusion criteria included: (1) child made a disclosure of sexual abuse and (2) caregiver provided consent for the data to be used for research.

2.1.2. Sample 2

Participant data for this sample were collected by research staff as part of a research study at a CAC in the southeastern United States. These assessments took place within six weeks of a forensic evaluation at the same CAC. Inclusion criteria included: (1) children were victims of child sexual abuse as determined by a professional forensic evaluation and (2) caregiver provided consent to participate in the research study.

Caregivers ($N = 386$) ranged in age from 23 to 75 years old ($M = 39.22$, $SD = 9.41$) and children were ages 6 to 17 years ($M = 12.13$, $SD = 2.72$; 39% school aged [6–11 years]; 61% adolescent [12–17 years]). Most caregivers in this sample identified as the child's mother (82%), were female (92%), and preferred to complete the assessment in English (82%) as opposed to Spanish (18%).

2.2. Measures

2.2.1. Maternal Self-report Support Questionnaire

All caregivers completed the 14-item Maternal Self-report Support Questionnaire (MSSQ; Smith et al., 2010), which consists of the 7-item emotional support subscale and the 7-item blame/doubt subscale. See Table 1 for the MSSQ items. Responses were made on a 7-point scale (0 = *not at all like me*, 1 = *kind of like me*, 2 = *a little like me*, 3 = *like me*, 4 = *quite a bit like me*, 5 = *a lot like me*, 6 = *very much like me*) and summed for each subscale. The Spanish-language version of the MSSQ was created following the International Test Commission (ITC) test development criteria (Hernández et al., 2020; Muñiz et al., 2013). The authors emphasized the need to ensure conceptual, linguistic, and metric equivalence across the English and Spanish versions. The adaptation was created by a team of bilingual professionals at the Sample 1 CAC through an iterative process of forward linguistic translation and subsequent backward translations. Several translators initially worked independently and subsequently formed a committee to review and compare the proposed translations, resolve discrepancies, and produce a consensus version. The bilingual translators consulted with the lead author to reduce conceptual discrepancies between the English- and Spanish-language versions and to ensure the instructions were clear and comprehensible to the target population. The initial Spanish translations were pilot tested and further modified by group consensus as needed to guarantee clarity and conceptual equivalence with the English version. In Sample 1, coefficient alpha was 0.80 for emotional support and 0.74 for blame/doubt. In Sample 2, coefficient alpha was 0.87 for emotional support and 0.72 for blame/doubt.

2.2.2. Demographics

2.2.2.1. Child age.
Caregivers completed a single item assessing child age.

2.2.2.2. Caregiver-child relationship.
Caregivers in Sample 1 answered a free response question assessing, "What is your relationship to the child?" Caregivers in Sample 2 answered a forced response question on their relationship to the child with 7 discrete categories (1 = *biological mother*, 2 = *adoptive mother*, 3 = *foster mother*, 4 = *stepmother*, 5 = *other biological relative*, 6 = *other*). The caregiver-child relationship was coded dichotomously across both samples such that mothers included any self-identified "mother" (e.g., *biological mother, stepmother, adoptive mother*) and non-mothers included all other caregiver-child relationships (e.g., *father, aunts, grandparents*).

2.2.2.3. Caregiver preferred language.
Caregivers answered a single item assessing their preferred language (i.e., English or Spanish).

2.3. Data analysis

To test our first hypothesis, we conducted Confirmatory Factor Analyses (CFA). All analyses were conducted with the software program R and its package lavaan (Rosseel, 2012). We examined the fit indices of the two-factor model assessing emotional support and blame/doubt. This model examined emotional support and blame/doubt as uncorrelated factors. We examined the Root Mean Square Error of Approximation (RMSEA), the 90% confidence interval for RMSEA, the Standardized Root Mean Square Residual (SRMR), and the Comparative Fit Index (CFI). According to Browne and Cudeck (1993) RMSEA values less than 0.05 indicate good fit,

Table 1
Standardized factor solutions.

Label	Item description	1	2
Emotional support			
ES1	Believed your child about everything that happened	0.506	
ES2	Willing to talk to your child about the abuse	0.559	
ES3	Often reassured your child that you would stand by her/him	0.669	
ES4	Tried to make your child feel safe	0.800	
ES5	Wanted to be supportive	0.739	
ES6	Told your child that s/he did the right thing by telling about the abuse	0.605	
ES7	Tried to be helpful	0.748	
Blame/doubt			
BD1	Could not help feeling angry with your child		0.431
BD2	Wondered if your child could have stopped the abuse if s/he wanted to		0.633
BD3	Questioned your child's honesty about the abuse		0.495
BD4	Wondered what your child might have done to stop the abuse from happening		0.743
BD5	Could not help resenting all the trouble your child's disclosure about the abuse caused		0.452
BD6	Told your child that s/he should have told you sooner about the abuse		0.498
BD7	Wondered if your child somehow brought the abuse on her/himself		0.184

Note: Standardized solutions are depicted. ES1–ES7 are items measuring emotional support; BD1–BD7 are items measuring blame/doubt; RMSEA (90% CI) = 0.07 (0.06, 0.08), SRMR = 0.06, CFI = 0.89.

and values between 0.05 and 0.08 indicate a reasonable fit. Hu and Bentler (1999) suggest that SRMR scores less than 0.08 indicate acceptable fit and both CFI scores greater than 0.90 indicate a reasonably good fit and scores greater than 0.95 indicate a good fit.

Inspection of the modification indices for the initial model suggested that allowing the error terms between 2 pairs of items on the emotional support scale (ES5 "*Wanted to be supportive*" and ES7 "*Tried to be helpful*;" and ES2 "*Willing to talk to your child about the abuse*" and ES6 "*Told your child that s/he did the right thing by telling about the abuse*") to covary would improve model fit. Therefore, we specified covariation of these two error terms and re-computed the model (Whittaker, 2012). In general, there is little consensus on the required sample size to conduct CFA, however, our present sample ($N = 386$) surpasses Tabachnick and Fidell's (2013) conservative suggestion of a minimum of 300 subjects.

To evaluate our second aim, we conducted Multigroup Confirmatory Factor Analyses (CFA) to examine the measurement invariance of the best-fitting factor model across groups. We conducted three group comparisons across: (1) child age (school age [6–11 years]; adolescent [12–17]); (2) caregiver-child relationship (mothers; non-mothers [fathers, aunts, grandparents, etc.]); and, (3) caregiver preferred language (English; Spanish). We used maximum likelihood estimation. Consistent with guidelines (e.g., Brown, 2015; Putnick & Bornstein, 2016), we examined measurement invariance by first testing the CFA model separately in each group. We then examined four different measurement invariance models: configural, metric, scalar, and residual. Each model requires stricter conditions to be met.

In the configural invariance model, all factor loadings and intercepts were allowed to load freely and differ across both groups (the least constrained model). Evidence of configural invariance indicates that both groups share the same underlying basic factor structure. In the metric invariance model, the factor loadings were constrained to be equal across both groups, but the intercepts were allowed to differ. Evidence of metric invariance indicates the latent factors have the same meaning across both groups but may have different units of measurement. In the scalar invariance model, the factor loadings and intercepts were constrained to be equal across both groups. Evidence of scalar invariance indicates the latent factors are related to the observed scores, allowing for mean-level comparisons across groups. In the residual invariance model, the factor loadings, intercepts, and indicator residual variances are constrained to be equal across groups (the most constrained model). Evidence of residual invariance indicates the item specific variance and error variance are similar across groups. Establishing residual variance is required for assessing full factorial invariance; however, it is not a prerequisite for testing mean-level differences between groups (Putnick & Bornstein, 2016).

To evaluate measurement invariance we conducted chi-square difference tests to examine whether the more restrictive models fit the data just as well as the less restrictive models. However, the chi-square difference test statistic is nearly always large and statistically significant among large samples and is sensitive to departures from multivariate normality. Therefore, we also compared the goodness of fit indices across each model by examining three alternative fit indices: RMSEA, SRMR, and CFI. RMSEA and CFI are included as standard evaluations of measurement invariance (Putnick & Bornstein, 2016). SRMR is included as it is a sensitive measure of loading invariance that is relatively independent of sample size (Chen, 2007; Putnick & Bornstein, 2016). Following Chen's (2007) recommendations, we interpreted acceptable model fit as: Δ RMSEA < 0.015, Δ SRMR < 0.03 for metric invariance, Δ SRMR < 0.015 for scalar or residual variance, and Δ CFI < |0.01|. For any comparisons that demonstrated noninvariance, we examined partial invariance to explore whether releasing parameter constraints would allow for mean-level comparisons across groups (Putnick & Bornstein, 2016).

For Multigroup CFA, it is generally recommended to have 100 participants per group (Wang & Wang, 2012). However, other widely accepted ratios include 5 to 10 participants per indicator variable (see Kyriazos, 2018, for review). Following these guidelines, our most disparate comparison (mothers vs. non-mothers) is adequate but may contribute to difficulties with model convergence and robust estimates.

Finally, we compared the individual item and mean levels of caregiver emotional support and blame/doubt across child age, caregiver-child relationship, and caregiver preferred language. Because many scientists and clinicians rely on observed scores (e.g., summed responses to a questionnaire) to test hypotheses and evaluate treatment progress, for groups that demonstrated measurement invariance, we conducted independent sample *t*-tests on the observed emotional support and blame/doubt scores. For any groups that demonstrated partial invariance, we compared latent factor scores. A power analysis of the most disparate comparison (mothers vs. non-mothers) using G*Power (Faul et al., 2007) indicated that with alpha set at 0.05, and our sample size of 386, power exceeded 0.98 to detect a medium-sized effect ($d = 0.50$).

3. Results

3.1. Confirmatory factor analysis

The standardized factor solutions for the two-factor model are reported in Table 1. The model demonstrated a reasonably good fit, RMSEA (90% CI) = 0.07 (0.06, 0.08), SRMR = 0.06, CFI = 0.89. According to Tabachnick and Fidell (2019), thresholds for evaluating standardized factor loadings are: 0.32 (poor), 0.45 (fair), 0.55 (good), and 0.71 (excellent). Factor loadings on the emotional support subscale ranged from 0.51 to 0.80, with six of the seven in the good to excellent range. Factor loadings on the blame/doubt subscale ranged from 0.18 to 0.74, with two of the seven in the good to excellent range and three in the fair range. The factor loading for item BD7 ("*Wondered if your child somehow brought the abuse on her/himself*") was in the poor range on the blame/doubt subscale. The two factors were weakly correlated, $r = -0.08$.

Given the reasonably good fit of the two-factor model, we proceeded with the examination of measurement invariance including all of the original MSSQ items, which is consistent with guidelines offered by Putnick and Bornstein (2016). These analyses are presented below. However, we also conducted an examination of measurement invariance excluding the item with poor factor loading (BD7) and

found the same pattern of results.

3.2. Measurement invariance

3.2.1. Child age

The measurement invariance of the two-factor structure across child age (school age, $n = 144$; adolescent, $n = 221$) was examined by multigroup CFA (see Table 2). The model demonstrated reasonably good fit for the school age group, RMSEA (90% CI) = 0.08 (0.06, 0.09), SRMR = 0.07, CFI = 0.92, and the adolescent group, RMSEA (90% CI) = 0.07 (0.06, 0.09), SRMR = 0.08, CFI = 0.88. Results of the chi-square difference test indicated no differences across the configural and metric invariance models, $^{\Delta}\chi^2 (13) = 15.42, p > .05$. However, the scalar invariance model fit significantly worse than the metric invariance model, $^{\Delta}\chi^2 (12) = 27.38, p < .05$. There was no difference between the scalar and residual invariance models, $^{\Delta}\chi^2 (2) = 2.61, p > .05$. Examination of the differences in fit indices (ΔRMSEA, ΔSRMR, and ΔCFI) supports full measurement invariance for the configural, metric, scalar, and residual models across child age.

3.2.2. Caregiver-child relationship

The results of the multigroup CFA across caregiver-child relationship (mothers, $n = 317$; non-mothers, $n = 69$) are reported in Table 2. The model demonstrated reasonably good fit across most fit indices for mothers, RMSEA (90% CI) = 0.07 (0.06, 0.08), SRMR = 0.07, CFI = 0.90, and non-mothers, RMSEA (90% CI) = 0.08 (0.03, 0.11), SRMR = 0.09, CFI = 0.93. The chi-square difference test suggested the metric invariance model demonstrated worse fit than the configural invariance model, $^{\Delta}\chi^2 (13) = 28.17, p < .05$. However, there was no difference in model fit between the metric and scalar invariance models, $^{\Delta}\chi^2 (12) = 12.99, p > .05$ or the scalar and residual invariance models, $^{\Delta}\chi^2 (2) = 0.53, p > .05$. Comparison of the fit indices (ΔRMSEA, ΔSRMR, and ΔCFI) supports full measurement invariance and indicated no difference in fit for the configural, metric, scalar, and residual models across caregiver-child relationship.

3.2.3. Caregiver preferred language

The results of the multigroup CFA across caregiver preferred language (English, $n = 316$; Spanish, $n = 70$) are reported in Table 2. The model demonstrated reasonably good fit across most fit indices for the English group, RMSEA (90% CI) = 0.08 (0.07, 0.09), SRMR = 0.07, CFI = 0.88. The model demonstrated poor fit for the Spanish group, RMSEA (90% CI) = 0.09 (0.05, 0.12), SRMR = 0.10, CFI = 0.86. The chi-square difference test and CFI fit index indicated the metric invariance model was a worse fit compared to the configural invariance model, $^{\Delta}\chi^2 (13) = 35.15, p < .001; \Delta$CFI = 0.02. Furthermore, the scalar invariance model emerged as worse fit compared to

Table 2
Model fit statistics across groups.

Model	χ^2 (df)	$^{\Delta}\chi^2$ (df)	RMSEA	ΔRMSEA	SRMR	ΔSRMR	CFI	ΔCFI
Child age								
School age (6–11 years, $n = 144$)	136.77 (73)		0.079		0.07		0.92	
Adolescent (12–17 years, $n = 221$)	160.50 (73)		0.074		0.08		0.88	
Configural	297.27 (146)		0.076		0.07		0.90	
Metric	312.68 (159)	15.42 (13)	0.074	0.002	0.08	0.01	0.90	0.00
Scalar	340.07 (171)	27.38 (12)*	0.075	0.001	0.08	0.00	0.89	−0.01
Residual	342.67 (173)	2.61 (2)	0.074	0.001	0.08	0.00	0.89	0.00
Caregiver-child relationship								
Mothers ($n = 317$)	184.10 (73)		0.070		0.07		0.90	
Non-mothers (fathers, aunts, etc., $n = 69$)	100.81 (73)		0.076		0.09		0.93	
Configural	384.91 (146)		0.071		0.07		0.91	
Metric	313.07 (159)	28.17 (13)*	0.072	0.001	0.08	0.01	0.90	−0.01
Scalar	326.06 (171)	12.99 (12)	0.069	0.002	0.08	0.00	0.89	−0.01
Residual	326.59 (173)	0.53 (2)	0.069	0.001	0.08	0.00	0.90	0.01
Caregiver preferred language								
English ($n = 316$)	219.83 (73)		0.081		0.07		0.88	
Spanish ($n = 70$)	109.94 (73)		0.085		0.10		0.86	
Configural	224.27 (146)		0.062		0.08		0.92	
Metric	259.42 (159)	35.15 (13)***	0.068	0.005	0.08	0.00	0.90	**−0.02**
Scalar	318.66 (171)	59.24 (12)***	0.079	0.011	0.10	**0.02**	0.85	**−0.05**
Residual	340.72 (173)	22.06 (2)***	0.084	0.005	0.08	**0.02**	0.83	**−0.02**
Partial scalar	270.79 (169)	11.37 (10)	0.066	0.002	0.09	0.01	0.90	0.00

Note: **Bold fit indices** represent significant differences in model fit. RMSEA: root mean square error of approximation; SRMR: standardized root mean square residual; CFI: comparative fit index. Configural: no constraints on loadings or intercepts between groups; Metric: loadings are constrained to be equal between groups; Scalar: loadings and intercepts are constrained to be equal between groups; Residual: loadings, intercepts, and indicator residual variances are constrained to be equal between groups; Partial Scalar: loadings and intercepts are constrained to be equal between groups except for the intercepts for two items "*Could not help resenting all the trouble your child's disclosure about the abuse caused*" and "*Often reassured your child that you would stand by her/him.*"

* $p < .05$.
*** $p < .001$.

the metric invariance model, $^{\Delta}\chi^2$ (12) = 59.24, $p < .001$, ΔSRMR = 0.02, ΔCFI = 0.05. Finally, the scalar model emerged as a worse fit compared to the residual invariance model, $^{\Delta}\chi^2$ (2) = 22.06, $p < .001$, ΔSRMR = 0.02, ΔCFI = 0.02. There were no differences in model fits across the RMSEA index. Overall, the comparison of models indicated measurement noninvariance across caregiver preferred language.

Given the evidence of noninvariance, consistent with guidelines from Putnick and Bornstein (2016) we examined the modification indices to assess whether releasing parameter constraints would allow for mean-level comparisons across caregiver preferred language. Examination of the modification indices suggested freeing the intercept of the blame/doubt item BD5 "*Could not help resenting all the trouble your child's disclosure about the abuse caused*" and the emotional support item ES3 "*Often reassured your child that you would stand by her/him*" would improve model fit. The results of the partial scalar invariance model in which all loadings and intercepts were constrained to be equal, except the intercepts of BD1 and ES3, are presented in Table 2. There was no difference in model fit between the metric and partial scalar invariance models, $^{\Delta}\chi^2$ (10) = 11.37, $p > .05$. Comparison of the fit indices supported partial measurement invariance across caregiver preferred language. This suggests that the measurement noninvariance is driven by the intercepts of BD1 and ES3.

3.2.4. Mean and item-level variance

Based on evidence of measurement invariance, we examined mean-level differences using observed scores for child age and caregiver-child relationship. Caregivers reported higher levels of blame/doubt towards adolescent children (aged 12 to 17 years; $M = 13.74$, $SD = 10.18$) compared to younger school-aged children (aged 6 to 11 years; $M = 11.72$, $SD = 8.83$), $t(361) = -2.01$, $p = .046$, $d = 0.21$. There were no differences in emotional support across child age ($p = .60$). There were no differences in emotional support or blame/doubt across caregiver-child relationship ($ps > .56$). Item-level differences across groups are presented in Table 3. Caregivers of adolescent children reported higher levels of blame/doubt on 2 of the 7 items. Specifically, BD1 "*Could not help feeling angry with your child*" and BD7 "*Wondered if your child somehow brought the abuse on her/himself.*" There were no item-level differences in emotional support across child age, or emotional support and blame/doubt across caregiver-child relationship.

We examined latent factor score differences for caregiver preferred language. Caregivers who completed the MSSQ in Spanish reported lower levels of emotional support compared to English-language caregivers, $d = 0.50$, $z = 4.92$, $p < .001$. Furthermore, Spanish-language caregivers reported higher levels of blame/doubt compared to English-language caregivers, $d = -0.40$, $z = -2.52$, $p = .012$. The item-level differences across caregiver preferred language are presented in Table 3. Because the invariance analyses suggested that scores on ES3 and BD1 represent nonequivalent units of measurement across caregiver preferred language mean comparisons were not computed for those items. Spanish-language caregivers reported lower levels of emotional support on every examined item compared to English-language caregivers, $ps < 0.05$. Spanish-language caregivers reported higher levels of blame/ doubt on two items, BD2 "*Wondered if your child could have stopped the abuse if s/he wanted to*" and BD5 "*Could not help resenting all the trouble your child's disclosure about the abuse caused,*" $ps < 0.05$.

Table 3
Item-level differences across groups.

	Item description	School age ($n = 144$)	Adolescent ($n = 221$)	Mothers ($n = 317$)	Non-mothers ($n = 69$)	English ($n = 316$)	Spanish ($n = 70$)
		Mean (SD)		Mean (SD)		Mean (SD)	
ES1	Believed your child about everything that happened	5.51 (1.08)	5.28 (1.37)	5.36 (1.25)	5.42 (1.29)	5.50 (1.15)	**5.14 (1.31)**
ES2	Willing to talk to your child about the abuse	5.47 (1.25)	5.28 (1.40)	5.33 (1.34)	5.42 (1.40)	5.41 (1.30)	**4.76 (1.70)**
ES3	Often reassured your child that you would stand by her/him	5.69 (0.93)	5.67 (1.05)	5.72 (0.84)	5.51 (1.45)	–	–
ES4	Tried to make your child feel safe	5.77 (0.73)	5.80 (0.70)	5.77 (0.70)	5.81 (0.73)	5.88 (0.50)	**5.37 (0.95)**
ES5	Wanted to be supportive	5.66 (0.93)	5.82 (0.62)	5.74 (0.79)	5.82 (0.60)	5.86 (0.56)	**5.30 (1.12)**
ES6	Told your child that s/he did the right thing by telling about the abuse	5.60 (1.22)	5.60 (1.17)	5.60 (1.15)	5.50 (1.43)	5.60 (1.26)	**5.21 (1.27)**
ES7	Tried to be helpful	5.65 (0.98)	5.61 (0.99)	5.60 (1.00)	5.72 (0.84)	5.76 (0.78)	**5.03 (1.33)**
BD1	Could not help feeling angry with your child	0.79 (1.83)	**1.33 (2.10)**	1.02 (1.92)	1.39 (2.18)	–	–
BD2	Wondered if your child could have stopped the abuse if s/he wanted to	1.43 (2.24)	1.70 (2.32)	1.62 (2.27)	1.90 (2.46)	1.43 (2.20)	**2.03 (2.28)**
BD3	Questioned your child's honesty about the abuse	1.94 (2.47)	2.38 (2.50)	2.23 (2.45)	2.04 (2.48)	2.10 (2.48)	2.07 (2.51)
BD4	Wondered what your child might have done to stop the abuse from happening	1.88 (2.31)	2.19 (2.34)	2.08 (2.28)	2.32 (2.62)	2.31 (2.32)	2.70 (2.51)
BD5	Could not help resenting all the trouble your child's disclosure about the abuse caused	2.07 (2.47)	1.89 (2.41)	1.98 (2.42)	1.99 (2.58)	1.57 (2.28)	**3.81 (2.33)**
BD6	Told your child that s/he should have told you sooner about the abuse	2.59 (2.64)	2.74 (2.66)	2.64 (2.64)	3.03 (2.71)	3.50 (2.60)	3.97 (2.26)
BD7	Wondered if your child somehow brought the abuse on her/himself	1.05 (2.07)	**1.51 (2.40)**	1.34 (2.28)	1.03 (2.10)	0.39 (1.25)	0.70 (1.70)

Note: ES1–ES7 are items measuring emotional support; BD1–BD7 are items measuring blame/doubt; **Boldface** values indicate independent sample *t*-test differences between groups at $p < .05$. The invariance analyses suggested that scores on ES3 and BD1 represent nonequivalent units of measurement across caregiver preferred language so mean comparisons were not computed.

4. Discussion

The present study examined the factor structure, measurement invariance, and mean-level differences of the Maternal Self-report Support Questionnaire (MSSQ) across a diverse group of caregivers and their children following a sexual abuse disclosure. Confirmatory factor analyses replicated the original two-factor solution, with the emotional support and blame/doubt subscales emerging as distinct factors. The model demonstrated reasonably good fit, but the final model included the covariance between two error terms and one factor loading was in the poor range. It seems probable that the covariance between the two error terms was driven by similarities in item content related to providing help (i.e., ES5 "*Wanted to be supportive*" and ES7 "*Tried to be helpful*") and talking to the child about the abuse (i.e., ES2 "*Willing to talk to your child about the abuse*" and ES6 "*Told your child that s/he did the right thing by telling about the abuse*"). Multigroup confirmatory factor analyses found the two-factor structure was invariant across child age and caregiver-child relationship, indicating that the MSSQ can be used to measure and meaningfully compare caregiver support following child sexual abuse disclosure across a range of child ages and caregiver-child relationships for English-speaking caregivers. However, results only supported partial scalar invariance across caregiver preferred language.

Although the MSSQ is arguably the most psychometrically sound published measure of caregiver support following child sexual abuse, one interpretation of the confirmatory factor analyses is that certain items on the MSSQ require further examination. Specifically, research on the MSSQ may benefit from additional exploration of the factor loadings on the blame/doubt subscale with larger samples. It is possible that characteristics unique to the present sample, such as the inclusion of non-mothers and Spanish-speaking caregivers, are driving differences in the observed factor structure compared to the initial validation of the MSSQ. In short, it is premature to suggest alterations are needed to the MSSQ items or subscales on the basis of the current confirmatory factor analyses. Nevertheless, the results do suggest a need for additional measurement research on the MSSQ, and caregiver support in general, following child sexual abuse.

The full measurement invariance found across child age indicates mean-level differences in caregiver support between school aged and adolescent children can be meaningfully interpreted. This suggests that researchers and clinicians can use the MSSQ to measure caregiver support across a wide age range of children and make direct comparisons of observed scores. The present study found that following sexual abuse disclosure, caregivers reported higher levels of blame/doubt towards adolescents compared to younger school-aged children. This is in line with previous research findings (Bolen & Lamb, 2002; Sufredini et al., 2020), and it is consistent with the idea that caregivers may think adolescents are capable of effective communication, which causes them to feel angry with their child or blame their child for failing to disclose the abuse sooner (Walsh et al., 2012). This may also reflect caregivers' appreciation of adolescents' agency in their sexual development and beliefs that adolescents are, in some way, responsible for the sexual abuse. One implication of this finding is that clinicians may consider using the MSSQ to measure differences in caregiver support across child age to inform the need for psychoeducation or interventions targeting caregiver's blame/doubt of their children.

We also found full measurement invariance across caregiver-child relationship indicating mean-level differences between mothers and non-mother caregivers in the present sample can be meaningfully interpreted. This indicates that researchers and clinicians can also use the MSSQ to measure caregiver support across mother and non-mother caregivers, combine the data from diverse types of caregivers, and make direct comparisons of observed scores. The present study found no differences in emotional support or blame/doubt across the caregiver-child relationship, indicating both groups of mothers and non-mothers responded in a similar manner across the MSSQ. These findings have important implications for future research examining differences between caregiver-child relationships in their association with caregiver support. For example, it is common in families affected by familial sexual abuse that the child is removed from the home and placed with a non-mother caregiver. The present findings suggest the MSSQ could still be used to measure caregiver support with the new non-mother caregiver. The present study suggests that the MSSQ might be an appropriate dependent measure in research seeking to improve caregiver support across different types of caregiver-child relationships. It may also be possible to examine whether caregiver support from distinct caregivers differentially affects child adjustment problems following sexual abuse. While future research would benefit from assessing additional diverse groups of caregiver-child relationships and examining combinations of caregiver-child relationships across child ages or genders, the present findings suggest the MSSQ can be used effectively across diverse groups of caregiver-child relationships.

We only found evidence of partial invariance across caregiver preferred language, suggesting, that although English- and Spanish-language versions of the MSSQ share the same underlying factor structure, there are certain items that have nonequivalent units of measurement. In general, there is scant research on the accuracy of mean-level tests for partially invariant models (Putnick & Bornstein, 2016). Vandenberg (2002) proposed that in order to confidently compare observed scores a "majority" of items should be invariant, and with only two of the 14 items displaying noninvariance, the present evaluation of the MSSQ meets this criterion. Nevertheless, additional research on the accuracy of partial invariance is needed and mean-level comparisons of observed scores across caregiver preferred language on the MSSQ should be done with caution. Across latent factor scores the present study found that Spanish-speaking caregivers reported lower levels of emotional support and higher levels of blame/doubt compared to English-speaking caregivers. Because evidence from nationally representative surveys indicates that Hispanic youth endorse higher levels of trauma exposure and greater rates of re-victimization compared to Caucasian youth (Andrews et al., 2015; Crouch et al., 2000), the ability to compare caregiver support following child sexual abuse between English- and Spanish-speaking caregivers represents an important, lingering gap in the literature.

Measurement noninvariance across preferred language can occur for many reasons. Despite efforts to achieve conceptual equivalence, there may be nuances in terms and concepts that contribute to differences in the degree to which English- and Spanish-speaking caregivers weigh their responses to certain items (Ortuño-Sierra et al., 2015). For example, a score of 3 on the item, "*Could not help resenting all the trouble your child's disclosure about the abuse caused*" for Spanish-speaking caregivers may have a very different meaning

than a score of 3 on this item for English-speaking caregivers, and are thus not comparable (White et al., 2011). Taken further, it may be the "trouble" caused by the child's abuse disclosure, such as law enforcement involvement, accessing medical exams, or therapy services, may have very different consequences and barriers for Spanish- vs English-speaking caregivers. In this way, cultural norms and systemic barriers can drive measurement noninvariance across preferred language. Nevertheless, interpretation of these differences across cultural norms does assume accurate linguistic translation and the Spanish-language version of the MSSQ has yet to undergo rigorous psychometric evaluation. Additional mixed-methods research involving Spanish-speaking caregivers is needed to examine the translation and validation of the Spanish-language version of the MSSQ.

The present study has several strengths, including one of the first examinations of measurement invariance in the assessment of caregiver support following child sexual abuse disclosure. Still there are several limitations. We did not obtain sufficient detailed information on caregiver demographic information. For example, we do not have information on the caregiver's country of origin, race, or ethnicity. There are several important cultural, social, and linguistic distinctions plausible within a group of Spanish-speaking caregivers, and it is possible that such differences could be driving the observed measurement noninvariance. Future research would benefit from examining measurement invariance across different caregiver countries of origin or comparing Spanish-versions developed for Central American countries and Puerto Rico. As noted above, there is also a limitation to interpreting mean-level differences in partial invariance models (Putnick & Bornstein, 2016), although latent factor scores can provide some preliminary evidence that these associations warrant further investigation.

Another limitation is that our caregiver-child relationship groups collapsed diverse, heterogenous types of caregiver-child relationships (e.g., fathers, aunts, grandparents). This limited our examination of measurement invariance to groups of mothers vs. non-mothers, as opposed to a more fine-grained analysis of different caregiver-child relationships. Although this is consistent with the current state of the literature on caregiver support following child sexual abuse disclosure (Bolen & Gergely, 2015), additional examinations across larger groups of caregivers, including fathers, grandparents, or kinship care placements is needed. Further, the present sample size of non-mothers ($n = 69$) prohibited our ability to examine comparisons between groups of caregiver-child relationships and child age (e.g., non-mothers of school aged children vs. non-mothers of adolescents). Relatedly, although adequate, a strong limitation is our sample size that resulted in Multigroup CFA comparisons with groups less than 100 people. Although we did not experience any difficulties with model convergence, the modest sample size can challenge our ability to obtain robust estimates. Finally, other psychometric characteristics of the MSSQ still need to be evaluated, including transient measurement error, to confidently determine the utility of this measure for diverse clinical and research purposes.

In conclusion, the present study extends the psychometric evaluation of the Maternal Self-report Support Questionnaire (MSSQ) and addresses a critical gap in our understanding of the measurement of caregiver support following child sexual abuse disclosure. We found adequate support for the original two-factor structure of the MSSQ, however, one item demonstrated poor loading, indicating a need for additional measurement research in this area. Although imperfect, the reasonable model fit of the MSSQ allowed us to proceed with examinations of measurement invariance. Establishing measurement invariance allows researchers and clinicians to draw confident conclusions regarding risk and protective factors or the differential effects of an intervention across diverse groups of caregivers and their children. The present findings showed measurement invariance across child age and caregiver-child relationships among English-speaking caregivers, suggesting the English version of the MSSQ can be used to compare observed scores in different age groups as well as diverse caregiver-child relationships. We found evidence of partial invariance for caregiver preferred language, indicating future research is still needed to establish observed score comparisons of caregiver support across English- and Spanish-speaking caregivers. Comparisons of mean-level and latent factor scores indicate caregiver support varies by child age and caregiver preferred language, respectively. Overall, findings suggest the MSSQ can be used to measure caregiver support across diverse groups of child ages and caregiver-child relationships, but caution should be practiced when interpreting mean differences in scores between English- and Spanish-speaking caregivers.

References

Andrews, A. R., Jobe-Shields, L., López, C. M., Metzger, I. W., De Arellano, M. A., Saunders, B., & Kilpatrick, D. G. (2015). Polyvictimization, income, and ethnic differences in trauma-related mental health during adolescence. *Social Psychiatry and Psychiatric Epidemiology, 50*(8), 1223–1234.
Bolen, R. M., & Gergely, K. B. (2015). A meta-analytic review of the relationship between nonoffending caregiver support and postdisclosure functioning in sexually abused children. *Trauma, Violence, & Abuse, 16*(3), 258–279.
Bolen, R. M., & Lamb, J. L. (2002). Guardian support of sexually abused children: A study of its predictors. *Child Maltreatment, 7*(3), 265–276.
Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford Publications.
Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Bollen (Eds.), *Testing Structural Equation Models* (pp. 136–162). Newbury Park, CA: Sage.
Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504.
Chen, F. F., & West, S. G. (2008). Measuring individualism and collectivism: The importance of considering differential components, reference groups, and measurement invariance. *Journal of Research in Personality, 42*(2), 259–294.
Cohen, J. A., & Mannarino, A. P. (2000). Predictors of treatment outcome in sexually abused children. *Child Abuse & Neglect, 24*(7), 983–994. https://doi.org/10.1016/S0145-2134(00)00153-8
Cohen, J. A., Mannarino, A. P., & Deblinger, E. (2017). *Treating trauma and traumatic grief in children and adolescents*. Guilford Publications.
<collab>U.S. Department of Commerce, C. B. (2019, August 20). Facts for features: Hispanic heritage month 2019 (Release No.: CB19-FF. 07). Retrieved from https://www.census.gov/newsroom/facts-for-features/2019/hispanic-heritage.html.
<collab>U.S. Department of Commerce, C. B. (2020, November 20). Stats for stories: National family caregivers month: November 2020 (Release No.: CB20-SFS.138). Retrieved from https://www.census.gov/newsroom/stories/family-caregivers-month.html.
Colman, R. A., Hardy, S. A., Albert, M., Raffaelli, M., & Crockett, L. (2006). Early predictors of self-regulation in middle childhood. *Infant and Child Development, 15*(4), 421–437.
Crouch, J. L., Hanson, R. F., Saunders, B. E., Kilpatrick, D. G., & Resnick, H. S. (2000). Income, race/ethnicity, and exposure to violence in youth: Results from the national survey of adolescents. *Journal of Community Psychology, 28*(6), 625–641.

Dunifon, R. (2013). The influence of grandparents on the lives of children and adolescents. *Child Development Perspectives, 7*(1), 55–60.

Elliott, A. N., & Carnes, C. N. (2001). Reactions of nonoffending parents to the sexual abuse of their child: A review of the literature. *Child Maltreatment, 6*(4), 314–331.

Everson, M. D., Hunter, W. M., & Runyan, D. K. (1989). *Parental reaction to abuse disclosure scale (PRADS)*. University of North Carolina.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Finkelhor, D., & Browne, A. (1985). The traumatic impact of child sexual abuse: A conceptualization. *American Journal of Orthopsychiatry, 55*(4), 530–541.

Hernández, B. A., Hidalgo, M. D., Hambleton, R. K., & Gómez-Benito, J. (2020). International test commission guidelines for test adaptation: A criterion checklist. *Psicothema, 32*(3), 390–398.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Jackson, S., & Goossens, L. (Eds.). (2020). *Handbook of adolescent development.* Psychology Press.

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology, 9*(08), 2207–2231.

Lewis, C., & Lamb, M. E. (2003). Fathers' influences on children's development: The evidence from two-parent families. *European Journal of Psychology of Education, 18* (2), 211–228.

Mannarino, A. P., & Cohen, J. A. (1996). Family-related variables and psychological symptom formation in sexually abused girls. *Journal of Child Sexual Abuse, 5*(1), 105–120.

Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). International test commission guidelines for test translation and adaptation. *Psicothema, 25*(2), 151–157.

Ortuño-Sierra, J., Chocarro, E., Fonseca-Pedrero, E., i Riba, S. S., & Muñiz, J. (2015). The assessment of emotional and behavioural problems: Internal structure of the Strengths and Difficulties Questionnaire. *International Journal of Clinical and Health Psychology, 15*(3), 265–273.

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.

Rollè, L., Gullotta, G., Trombetta, T., Curti, L., Gerino, E., Brustia, P., & Caldarera, A. M. (2019). Father involvement and cognitive development in early and middle childhood: A systematic review. *Frontiers in Psychology, 10*, 2405.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software, 48*(2), 1–36.

Rothenberg, W. A., Lansford, J. E., Alampay, L. P., Al-Hassan, S. M., Bacchini, D., Bornstein, M. H.Yotanyamaneewong, S., … (2020). Examining effects of mother and father warmth and control on child externalizing and internalizing problems from age 8 to 13 in nine countries. *Development and Psychopathology, 32*(3), 1113–1137.

Sangawi, H., Adams, J., & Reissland, N. (2015). The effects of parenting styles on behavioral problems in primary school children: A cross-cultural review. *Asian Social Science, 11*(22), 171–186.

Smith, D. W., Sawyer, G. K., Jones, L. M., Cross, T., McCart, M. R., & Ralston, M. E. (2010). Mother reports of maternal support following child sexual abuse: Preliminary psychometric data on the Maternal Self-report Support Questionnaire (MSSQ). *Child Abuse & Neglect, 34*(10), 784–792.

Spaccarelli, S. (1994). Stress, appraisal, and coping in child sexual abuse: A theoretical and empirical review. *Psychological Bulletin, 116*(2), 340–362.

Sufredini, F., Moré, C. L. O. O., Krenkel, S., & Crepaldi, M. A. (2020). Narratives of mothers whose children had been sexually abused: Maternal reactions and comprehension regarding child and adolescent sexual abuse. *Journal of Interpersonal Violence, 0*(0), 1–26. https://doi.org/10.1177/0886260520948520

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics: International edition.* Pearson.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (Seventh). Pearson.

Thomas, R., & Zimmer-Gembeck, M. J. (2012). Parent-Child Interaction Therapy: An evidence-based treatment for child maltreatment. *Child Maltreatment, 17*(3), 253–266.

Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods, 5*, 139–158. https://doi.org/10.1177/1094428102005002001

Walsh, W. A., Cross, T. P., & Jones, L. M. (2012). Do parents blame or doubt their child more when sexually abused by adolescents versus adults? *Journal of Interpersonal Violence, 27*(3), 453–470.

Wamser-Nanney, R. (2017). Maternal support following childhood sexual abuse: Links to parent-reported children's outcomes. *Child Abuse & Neglect, 67*, 44–53.

Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus.* Wiley, Higher Education Press.

White, R. M., Umaña-Taylor, A. J., Knight, G. P., & Zeiders, K. H. (2011). Language measurement equivalence of the Ethnic Identity Scale with Mexican American early adolescents. *The Journal of Early Adolescence, 31*(6), 817–852.

Whittaker, T. A. (2012). Using the modification index and standardized expected parameter change for model modification. *The Journal of Experimental Education, 80* (1), 26–44.

Winokur, M. A., Holtan, A., & Batchelder, K. E. (2018). Systematic review of kinship care effects on safety, permanency, and well-being outcomes. *Research on Social Work Practice, 28*(1), 19–32.