Title

Revisiting the Structural Validity of the Childhood Trauma Questionnaire (CTQ-SF) among Brazilian Adolescents: Insights from a Factor and Network Analysis.

Author Listing:

Michael Eduardo Reichenheim[1], PhD, Antonio Piolanti[2], PhD, Claudia Leite de Moraes[1,3], PhD.

Affiliation

[1] Department of Epidemiology, Institute of Social Medicine, Rio de Janeiro State University, Brazil.

[2] Health Psychology Unit, Institute of Psychology, University of Klagenfurt, Austria.

[3] Medical School, Estácio de Sá University, Rio de Janeiro, Brazil.

Publication Status: Not yet submitted for publication

Draft Date: 04.10.2024

Corresponding author

Antonio Piolanti

Health Psychology Unit, Institute of Psychology, University of Klagenfurt, Universitätsstraße 65-67, 9020 Klagenfurt, Austria

E-mail: antonio.piolanti@aau.at

Telephone: +43 463 2700 1676

Abstract

Child maltreatment is a serious global health issue, and the Childhood Trauma Questionnaire (CTQ-SF) is widely used to assess child abuse. However, its psychometric properties among adolescents have yet to be thoroughly explored. This study examined the factorial and network structure of the CTQ-SF among 721 Brazilian adolescents aged 15-19. Initial confirmatory factor analysis (CFA) assessed the original 25-item, five-factor solution, followed by exploratory models that revealed mis-specifications. Refinement required removing items from the physical neglect factor, resulting in a 20-item, four-factor model covering emotional neglect, physical abuse, sexual abuse, and emotional abuse. This solution demonstrated factorial item pertinence, unambiguity and reliability, and factor-based convergent and discriminant validity. The network analysis further highlighted strong links between emotional and physical abuse. The findings suggest that while the original model did not fully fit this population, the revised 20-item version offers improved and promising psychometric properties for assessing childhood trauma in adolescents.

Keywords: Childhood Trauma; CTQ-SF; Psychometric Analysis; Network Analysis; Adolescents; Brazil.

1. Introduction

Child maltreatment is a significant clinical and public health concern with extensive and long-lasting health impacts. In addition to chronic physical conditions such as obesity, cardiovascular problems, diabetes and autoimmune diseases, trauma and abuse during early childhood have been consistently associated with a range of psychiatric and related disorders (Buss et al., 2017; Norman et al., 2012). The literature reports that children exposed to maltreatment are at increased risk for substance abuse, depression and anxiety, post-traumatic stress disorder (PTSD) and suicidal behavior later in life (Buss et al., 2017; Norman et al., 2012).

Given its profound impacts, child maltreatment is regarded as one of the most important preventable factors for poor health outcomes (The Lancet Public Health, 2021). In response, target 16.2 of the 2030 UN Agenda for Sustainable Development Goals calls for the elimination of all forms of violence against children (United Nations, 2015). Achieving this objective necessitates the development of tailored and effective prevention programs and policies, which depend on reliable and valid measurement tools to identify all types of child abuse and neglect. Among these tools, the Childhood Trauma Questionnaire-CTQ stands out as one of the most widely used measures (Meinck et al., 2022).

The CTQ is a retrospective self-report instrument developed by Bernstein and colleagues in 1994 in the United States. The initial version comprised 70 items rated on a 5-point scale from 1 (*never true*) to 5 (*very often true*) and organized into five dimensions. Physical neglect (PN)

purportedly assesses caregivers' failure to provide physical needs such as clothing, food, and safety. The emotional neglect (EN) subscale evaluates the lack of provision for the child's fundamental emotional needs. The physical abuse (PA) subscale appraises experiences of bodily assaults by an older individual. The sexual abuse (SA) subscale measures sexual behaviors towards a child by an older person. Lastly, the emotional abuse (EA) subscale gauges childhood experiences of verbal abuse, such as humiliation and harassment directed at the child's self-esteem and well-being by an adult or older individual (Bernstein et al., 1994).

A few years after its development, the CTQ was shortened to improve its applicability. Specifically, the developers implemented an Exploratory Factor Analysis and retained items with factor loadings above 0.50 on their intended factors and below 0.30 on the remaining factors. This process resulted in a streamlined set of 25 questions, with 5 items *per* original dimension. The validity of this new format was corroborated through Confirmatory Factor Analysis (CFA) and measurement invariance was observed across adolescents and adults (Bernstein et al., 2003). The abbreviated version was named CTQ-SF, where "SF" stands for *short form*. Since its development, the CTQ-SF has been translated into several languages, including French (Paquette et al., 2004), Brazilian Portuguese (Grassi-Oliveira et al., 2006), Persian (Garrusi & Nakhaee, 2009), Dutch (Thombs et al., 2009), German (Wingenfeld et al., 2010), Italian (Innamorati et al., 2016), Slovak (Petrikova et al., 2021), Spanish (Hernandez et al., 2013), and Chinese (Wang et al., 2022).

The psychometric properties of the CTQ-SF have predominantly been investigated in adult samples, with limited research focusing on adolescents. However, adolescence is a critical period of development where experiences of maltreatment can negatively affect neuro-biological and psychological systems (Engstrom et al., 2021), school performance (Romano et al., 2015) as well as the response to psychotherapy (Waldron et al., 2019). Moreover, child maltreatment increases the risk of alcohol and illicit drug use during adolescence, as well as teen dating violence (Richards et al., 2017; Tussey et al., 2018; Varlioglu & Hayes, 2022). Therefore, accurately assessing child maltreatment in adolescence is crucial for research, clinical practice and health policy formulation.

An inspection of the psychometric history of the CTQ-SF among adolescents reveals that internal consistency has been the most frequently assessed property. Estimates using Cronbach's α or McDonald's ω coefficients have ranged from 0.33 to 0.79 for PN, 0.71 to 0.90 for EN, 0.68 to 0.90 for PA, 0.64 to 0.96 for SA, and 0.67 to 0.91 for EA (Aloba et al., 2020; Bernstein et al., 2003; Charak et al., 2017; Dovran et al., 2013; Hagborg et al., 2022; García-Fernández et al., 2024; Mizuki et al., 2021; Peng et al., 2023; Wang et al., 2022). One study evaluated the four-week interval test-retest reliability employing Intraclass Correlation Coefficients (ICC), and reported estimates of 0.59, 0.70, 0.54, 0.63 and 0.44 for the respective subscales (Wang et al., 2022). Furthermore, findings support the construct validity of the CTQ-SF among adolescents. Positive correlations were found between the CTQ-SF subscales and measures evaluating psychological distress, such as the

General Health Questionnaire, Beck-Depression Inventory, Patient Health Questionnaire, the State Trait Anxiety Inventory, the Hospital Anxiety and Depression Scale, Non-Suicidal Self-Injury, and the Positive and Negative Suicide Ideation Inventory (Aloba et al., 2020; García-Fernández et al., 2024; Wang et al., 2022). Conversely, negative correlations were shown with measures of family cohesion, perceived parental support and psychological health (Hagborg et al., 2022).

In addition to the factors analyses conducted by Bernstein and colleagues (2003), seven studies have explored the configural structure of the CTQ-SF among adolescents, all via CFA, yielding mixed results (Aloba et al., 2020; Charak et al., 2017; Dovran et al., 2013; García-Fernández et al., 2024; Grassi-Oliveira et al., 2014; Peng et al., 2023; Wang et al., 2022). Three studies reported a structure consistent with the original model, both regarding the number of factors and the placement of items within their theorized dimension (Dovran et al., 2013; García-Fernández et al., 2024; Wang et al., 2022). Conversely, one study suggested the removal of an item from the SA dimension (Charak et al., 2017), while two studies reported a model where two items originally postulated to belong to the PN factor were part of the EN dimension (Aloba et al., 2020; Grassi-Oliveira et al., 2014). The remaining study removed one item pertaining to PN, and proposed a 4-factor solution where PN and EN were combined into a single neglect dimension (Peng et al., 2023).

Among the six studies mentioned, only one assessed item ambiguity through the presence of cross-loadings, and item redundancy, expressed by residual correlations (Aloba et al., 2020). The proposed final model

included residual correlations between a pair of items in the PA factor and another in the SA factor. Notably, residual correlations were also reported at the inception of the CTQ development process in 2003 but were not addressed at the time (Berstein, 2003).

Additionally, there are some psychometric properties, such as factor-based convergent and discriminant validities, that have never been examined by previous studies. Convergent validity assesses whether items share more information within their assigned factor than outside the system through measurement errors, while discriminant factorial validity indicates whether there is more information internal to the component factors than across them (Reichenheim & Bastos, 2021). Several studies on the CTQ-SF have reported relatively high factorial correlations (0.83-0.99) yet never explicitly acknowledging any potential violations of discriminant factorial validity (Aloba et al., 2020; Bernstein et al., 2003; Charak et al., 2017; García-Fernández et al., 2024). Finally, it is important to note that no previous study has conducted a graphical network analysis of the CTQ-SF among adolescents to assess its psychometric properties. This approach may provide novel insights into the connectivity of the items and dimensions of measurement tools (Epskamp et al., 2018).

Evidence from the current state of the art highlights that several important issues regarding the application of the CTQ-SF to adolescents remain to be studied. One is whether the five-tiered dimensional structure is sustainable in this population. Another is whether item-factor ambiguities (cross-loadings) and content redundancies (residual

correlations) will replicate as observed in previous studies. A third issue concerns the corroboration of factor-based convergent and discriminant validities. The main objective of this study is thus to revisit and further explore these configural and metric properties among Brazilian adolescents. As an ancillary objective, we also provide insight into how items relate within and across factors using exploratory graphical network analyses.

2. Methods

2.1. Sampling and participants

This was a cross-sectional study with data collected from October 2016 to February 2017. Participants were second-year high school students attending public and private institutions of the IX Administrative Region (IX-AR) of Rio de Janeiro, Brazil. The IX-AR is characterized by high population density, including approximately 190,000 people from various socioeconomic backgrounds residing in areas ranging from fully urbanized neighborhoods to slums (favelas) with inadequate access to urban services and infrastructure (Instituto Pereira Passos, 2020).

In 2016, the region had five public and fifteen private schools, comprising 29 and 23 secondary school classes with 714 and 756 students, respectively. Participants were selected using multi-stage stratified sampling. A total of 26 classes from thirteen schools (749 students) were selected randomly with probability proportional to the school sizes. Additional sampling was conducted within this group to

ensure sufficient public-school students and participant weights were adjusted to reflect the number of eligible individuals. This procedure resulted in varying sampling weights for students from different schools. All students within the selected classes were invited to participate in the study. Only 28 declined, resulting in a participation rate of 96.3% (721 students).

2.2. Measurement tools and procedures

This study employed the QUESI (*Questionário Sobre Traumas na Infância*), a Brazilian Portuguese adaptation of the CTQ-SF (Grassi-Oliveira et al., 2006). As mentioned, the five-dimensional instrument includes 25 component items, each rated on a five-point scale, resulting in total subscale scores ranging from 5 to 25. The recall period for the experienced traumas focused on the first ten years of the adolescents' lives.

The assessment form in this study also included modules concerning demographic, economic, family, and school-related information and was administered to students in their classrooms while supervised by a fieldwork coordinator.

2.3. Data analysis

Data analysis was conducted in six steps. In Step 1, the model proposed by Bernstein and colleagues (2003) was tested via CFA (Brown, 2015). The strength of the factor loadings and their respective residuals were used as empirical representatives of the items' reliability (discriminating ability). Items with loadings < 0.40 or residual variance > 0.70 were

regarded as insufficiently reliable and thus inadmissible (Ford, MacCallum & Tait, 1986). Modification Indices (MI; univariate Lagrange multiplier) and Expected Parameter Changes (EPC) (Brown, 2015; Sorbom, 1989) were evaluated to check for potential cross-loadings and residual correlations. MI indices indicate the reduction in the chi-square value of the model when a specific parameter is added, while EPC indices estimate the presumed absolute magnitude of the change in the parameter if later estimated freely (Brown, 2015). To identify potential residual correlations, we initially focused on unusually large MIs and EPCs with yx-standardized correlations > 0.30. Residual correlations were freely estimated if this threshold was met for a given pair of items. A corroborated value > 0.30 served as an indication of conditional non-independence (Reeve et al., 2007).

Given the issues that emerged in the previous step (as detailed in the Results), an exploratory approach was subsequently undertaken (Step 2). Horn's Parallel Analysis (HPA) using principal component analysis (PCA) was first implemented to explore how many factors could be retained, assessing the number of adjusted eigenvalues above 1.0 (Dinno, 2009; Gorsuch, 1983; Horn, 1965). The progression of eigenvalues in the scree plots was also examined as a supplementary procedure. An Exploratory Structural Equation Model (ESEM) using Geomin oblique rotation was conducted thereafter to further explore the factor structure, including possible cross-loadings and residual correlations (Asparouhov & Muthen, 2009). Steps 3 and 4 involved a sequence of new exploratory models (HPA and related ESEM) with different sets of items to further identify

factor or item mis-specifications. In step 5, we reverted to a confirmatory framework by applying a CFA to the best-fitting solution achieved earlier. The strength of the factor loadings and their respective residuals, potential cross-loadings and residual correlations were then re-evaluated. The magnitude of the factorial correlations was also assessed, regarding values of 0.85 or above as highly suspicious of violating factor-based discriminant validity (FDV) (Brown, 2015). Regardless, the FDV was statistically evaluated by testing the difference between the square root of the Average Variance Extracted (AVE) of each factor and all correlations involving this factor. The AVE ranges from 0 to 1, and measures the amount of variance captured by a factor in relation to the amount of variance due to measurement error. If there is FDV, more information is expected to "flow" from each factor to the items than between the factors themselves (Reichenheim & Bastos, 2021). Thus, FDV of any factor k would be rejected when $\sqrt{AVE_{(f_k)}} - \Phi_{(k', f_{\neq k})}$ was negative and statistically significant. Any other result would support its corroboration. Factor-based convergent validity (FCV) would be endorsed if the AVE ≥ 0.5, indicating that at least 50% of the variance is due to the underlying factor rather than the joint information contained in the item errors (Brown, 2015; Hair et al., 2010). For completeness, McDonald's ω (omega) reliability coefficients were also computed (McDonald, 1999).

All five steps involving factor analyses employed probit models on polychoric matrices, using the robust diagonally weighted least squares estimator (WLSMV). This estimator also allows for accommodating missing data (Finney & Distefano, 2013). Model fits were evaluated with

the root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis index (TLI) (Brown, 2015). RMSEA values < 0.06 indicates good fit of the model, while > 0.10 suggests inadequate fit and model rejection. CFI and TLI values vary from 0 to 1, with > 0.95 indicating an acceptable fit.

Lastly, in step 6, an Exploratory Graphical Analysis (EGA) was conducted on the items identified in the final model (Golino & Epskamp, 2017). This analysis had three main goals: (1) to uncover the emerging item clustering pattern by using a graphical model-based procedure; (2) to assess the strength of the links between the items positioned in each identified graphical cluster (known as *community* in the parlance of graphical networks) and (3) to assess the strength of the links between items located in different clusters.

To this end, we employed a network approach based on the Gaussian Graphical Model (Lauritzen, 1996). In this network structure, items are represented by nodes and each edge represents the partial correlation coefficient between two nodes while conditioning on the other variables in the network (Golino & Epskamp, 2017). The strength of the relationship between nodes is indicated in terms of edge weights (diagrammatically reflected by the thickness of the edge). To estimate the network, we used the EGAnet R-package and applied the EGA function (Golino & Christensen, 2020) to the polychoric correlation matrix previously generated in Mplus. The function uses the graphical least absolute shrinkage and selection operator (glasso) algorithm, which involves a regularization technique that shrinks trivial small coefficients

to zero (Golino & Epskamp, 2017). This leads to a parsimonious network in which only a few edges are included to explain the associations. The EGA function also applies the walktrap algorithm to detect communities of highly connected nodes (equivalent to the latent factor) (Golino & Epskamp, 2017).

Descriptive analyses were conducted in Stata version 16.1 (StataCorp, 2019). All factorial analyses were executed in Mplus version 8.11 (Muthén & Muthén, 1998-2023). The network analysis was performed in R version 4.3.3 (R Core Team, 2024) via Studio 2023.12.1 Build 402 interface (RStudio Team, 2024).

3. Results

3.1. Socio-demographic characteristics of the study sample

The mean age of participants was 17.4 years, with a nearly equal distribution between male and female respondents. Approximately 32% identified as brown/mixed or mestizos and 16% as black; 42.5% were living with both parents, and 27.7% were from the lower socioeconomic classes. About 40% of the adolescents were enrolled in public schools, and 10% were attending evening classes.

3.2. Structural Validity of the CTQ-SF

Table 1 shows the results of the CFA aimed at corroborating the original five-factor model proposed by Bernstein and colleagues (2003).

Table 1. Dimensional structure of the original CTQ-SF: Confirmatory Factor Analysis of the 25-item, 5-factor solution.

	Items	$\lambda_i^{a}{}_{(f\,1^{c})}$	$\lambda_i^{a}{}_{(f\,2^{d})}$	$\lambda_{i(f\,3^{e})}$	$\lambda_i^{a}{}_{(f\,4^{f})}$	$\lambda_i^{a}{}_{(f\,5^{g})}$	δ_i^{b}
i1.	*Not enough to eat*	.261					.932
i2.	*Got (not) taken care of*	**.561**					.685
i4.	*Parents were high or drunk*	**.451**					.796
i6.	*Wore dirty clothes*	**.404**					.836
i26.	*Got (not) taken to a doctor*	**.563**					.683
i5.	*(Not) made me feel important*		**.609**				.630
i7.	*(Not) felt loved*		**.783**				.386
i13.	*Was (not) looked out for*		**.811**				.342
i19.	*Family (not) felt close*		**.805**				.352
i28.	*Family was (not) source of strength*		**.846**				.285
i9.	*Hit hard enough to see doctor*			**.854**			.271
i11.	*Hit hard enough to leave bruises*			**.855**			.269
i12.	*Punished with hard objects*			**.668**			.554
i15.	*Was physically abused*			**.865**			.253
i17.	*Hit badly enough to be noticed*			**.863**			.255
i20.	*Was touched sexually*				**.903**		.184
i21.	*Hurt if I did not do something sexual*				**.798**		.363
i23.	*Made me do sexual things*				**.848**		.280
i24.	*Was molested*				**.965**		.069
i27.	*Was sexually abused*				**.923**		.148
i3.	*Called names by family*					**.653**	.573
i8.	*Parents wished I was never born*					**.730**	.467

i14.	*Family said hurtful things*	**.834**	.304
i18.	*Felt hated by family*	**.727**	.472
i25.	*Was emotionally abused*	**.817**	.333
	RMSEA [h]	.020 (.013; .026)	
	CFI [i]	.985	
	TLI [l]	.984	

Notes. [a] Loadings (Standardized). [b] Measurement Errors (Uniqueness). [c] Physical Neglect. [d] Emotional Neglect. [e] Physical Abuse. [f] Sexual Abuse. [g] Emotional Abuse. [h] RMSEA = Root Mean Square Error of Approximation; In brackets: 90% confidence intervals. [i] CFI = Comparative Fit Index. [l] TLI = Tucker-Lewis Index.

The RMSEA, CFI, and TLI indices suggested adequate fit. Item i1 showed a loading < 0.40 and, along with items i4 and i6, a residual variance > 0.70; item i2 and i26 showed residual variances of approximately 0.70. Furthermore, diagnostics suggested a possible residual correlation between items i13 and i19 (MI = 18.248 / EPC = 0.620), which yielded a free estimate of $\rho_{i13 \leftrightarrow i19} = 0.452$.

In light of these results, an exploratory framework was adopted (Step 2). The HPA on all 25 items indicated five eigenvalues above 1.0. Table 2 presents the results of the complementary ESEM.

Table 2. Dimensional structure of the original CTQ-SF: Exploratory Factor Analysis of the 5-factor solution.

	Items	$\lambda_i^{a}{}_{(f1^{c})}$	$\lambda_i^{a}{}_{(f2^{d})}$	$\lambda_i{}_{(f3^{e})}$	$\lambda_i^{a}{}_{(f4^{f})}$	$\lambda_i^{a}{}_{(f5^{g})}$	δ_i^{b}
i1.	*Not enough to eat*	.041	.135	.121	.079	–.022	0.939
i2.	*Got (not) taken care of*	**.534**	.054	.085	.184	–.036	0.592
i4.	*Parents were high or drunk*	.069	.210	.103	.077	.104	0.840
i6.	*Wore dirty clothes*	–.015	.122	.319	.011	.046	0.840
i26.	*Got (not) taken to a doctor*	.329	.023	–.042	.124	.250	0.722
i5.	*(Not) made me feel important*	**.725**	.188	.001	–.031	–.002	0.364
i7.	*(Not) felt loved*	**.508**	.257	.032	–.009	.245	0.362

i13.	*Was (not) looked out for*	.229	**.716**	.036	.014	.019	0.286
i19.	*Family (not) felt close*	.020	**.810**	.003	–.044	.180	0.171
i28.	*Family was (not) source of strength*	.301	**.420**	–.031	.115	.284	0.337
i9.	*Hit hard enough to see doctor*	–.009	.084	**.896**	.143	–.174	0.214
i11.	*Hit hard enough to leave bruises*	.073	–.171	**.778**	–.014	.170	0.272
i12.	*Punished with hard objects*	–.047	–.130	**.686**	.000	.102	0.508
i15.	*Was physically abused*	.089	–.002	**.794**	–.027	.050	0.290
i17.	*Hit badly enough to be noticed*	–.075	.022	**.907**	–.016	.005	0.202
i20.	*Was touched sexually*	–.240	.151	–.012	**.850**	.197	0.094
i21.	*Hurt if I did not do something sexual*	.290	–.207	–.045	**.764**	.076	0.285
i23.	*Made me do sexual things*	.145	–.133	.019	**.862**	–.027	0.240
i24.	*Was molested*	–.013	.031	.089	**.933**	–.007	0.068
i27.	*Was sexually abused*	.003	.002	.173	**.894**	–.073	0.136
i3.	*Called names by family*	.052	.012	.181	–.008	**.513**	0.577
i8.	*Parents wished I was never born*	.228	–.007	.054	–.142	**.671**	0.414
i14.	*Family said hurtful things*	–.046	.001	.044	.138	**.824**	0.381
i18.	*Felt hated by family*	–.064	.135	–.018	.025	**.706**	0.204
i25.	*Was emotionally abused*	.098	.056	.194	.072	**.563**	0.421
RMSEA [h]	.016 (.004; .024)						
CFI [i]	.993						
TLI [l]	.989						

Notes. [a] Loadings (Standardized). [b] Measurement Errors (Uniqueness). [c] Physical Neglect. [d] Emotional Neglect. [e] Physical Abuse. [f] Sexual Abuse. [g] Emotional Abuse. [h] RMSEA = Root Mean Square Error of Approximation; In brackets: 90% confidence intervals. [i] CFI = Comparative Fit Index. [l] TLI = Tucker-Lewis Index.

Although a five-factor structure re-emerged (as indicated by the fit indexes), only three factors (3, 4 and 5) were consistent with the respective dimensions of PA, SA and EA postulated by Bernstein and

colleagues (2003). Four items (i1, i4, i6 and i26) theorized to belong to the PN factor showed loadings < 0.40 across all factors and residuals > 0.70. Moreover, two items (i5 and i7) originally conjectured as part of the EN factor, joined one item (i2) from PN into a separate factor. The remaining three items (i13, i19 and i28) from EN formed a single factor. The ESEM also suggested a possible residual correlation between items i11 and i12 of the PA factor (MI = 10.114 / EPC = 0.539; $\rho_{i11 \leftrightarrow i12} = 0.393$).

In Step 3, we removed the problematic items identified previously (i1, i4, i6 and i26) and re-applied an exploratory analysis to the reduced 21-item model. The HPA indicated four eigenvalues above 1.0. The related four-factor ESEM (Supplementary Material, Table S1), showed that only factor 4 (SA) was consistent with the structure originally postulated (Bernstein, 2003). Item i2, originally postulated to be part of the PN factor, formed a single factor with the EN items, namely, i5, i7, i13, i19 and i23. The latter three also cross-loaded in the fourth extracted factor primarily holding the postulated EA items. Interim diagnostics showed two potential residual correlations involving pairs of items i11 $\leftrightarrow$ i12 (MI = 17.064 / EPC = 0.559) and i13 $\leftrightarrow$ i19 (MI = 36.877 / EPC = 0.946). The free estimations were $\rho_{i11 \leftrightarrow i12} = 0.404$ and $\rho_{i13 \leftrightarrow i19} = 0.515$, respectively.

We removed item i2 from the model, as it was not originally theorized to be part of the EN factor and reverted to an exploratory approach on the remaining 20 items (Step 4). The HPA showed only three factors with adjusted eigenvalues strictly above 1.0. However, the related scree plot (Supplementary Material, Figure S1) showed eigenvalues levelling off

only beyond the fourth eigenvalue. In accordance with the original configuration (once excluding the PN items), we retained and further explored a four-factor structure. The related ESEM (Supplementary Material, Table S2) fitted well, and item pertinence was also consistent. Loadings and residual variances were within acceptable values, and no cross-loadings were observed. The interim diagnostics continued to indicate the presence of the two residual correlations, involving the pairs i11 ↔ i12 (MI = 16.637 and EPC = 0.534) and i13 ↔ i19 (MI = 34.743 and EPC = 0.888). The free estimates were $\rho_{i11 \leftrightarrow i12} = 0.396$ and $\rho_{i13 \leftrightarrow i19} = 0.489$, respectively.

Table 3 presents the results of the CFA for the 4-factor solution applied to the 20 items (Step 5).

Table 3. Dimensional structure of the modified 20-item CTQ-SF: Confirmatory Factor Analysis of the 4-factor solution.

	Items	λ_i^a(f 1[c])	λ_i^a(f 2[d])	λ_i^a(f 3[e])	λ_i^a(f 4[f])	δ_i^b
i5.	(Not) made me feel important	**.594**				.647
i7.	(Not) felt loved	**.778**				.395
i13.	Was (not) looked out for	**.805**				.351
i19.	Family (not) felt close	**.818**				.332
i28.	Family was (not) source of strength	**.847**				.282
i9.	Hit hard enough to see doctor		**.847**			.283
i11.	Hit hard enough to leave bruises		**.857**			.265
i12.	Punished with hard objects		**.669**			.553
i15.	Was physically abused		**.864**			.253
i17.	Hit badly enough to be noticed		**.868**			.247
i20.	Was touched sexually			**.906**		.179

i21.	Hurt if I did not do something sexual	**.795**		.368
i23.	Made me do sexual things	**.849**		.280
i24.	Was molested	**.965**		.069
i27.	Was sexually abused	**.922**		.150
i3.	Called names by family		**.654**	.573
i8.	Parents wished I was never born		**.725**	.475
i14.	Family said hurtful things		**.841**	.293
i18.	Felt hated by family		**.726**	.473
i25.	Was emotionally abused		**.813**	.339
	RMSEA [g]	.025 (.018; .032)		
	CFI [h]	.987		
	TLI [i]	.985		

Notes. [a] Loadings (Standardized). [b] Measurement Errors (Uniqueness). [c] Emotional Neglect. [d] Physical Abuse. [e] Sexual Abuse. [f] Emotional Abuse. [g] RMSEA = Root Mean Square Error of Approximation; In brackets: 90% confidence intervals. [h] CFI = Comparative Fit Index. [i] TLI = Tucker-Lewis Index.

The model adjusted adequately, loadings ranged from 0.594 to 0.965, and there were no items presenting residuals > 0.7. The MI (22.810) and EPC (0.635) remained indicative of a $i13 \leftrightarrow i19$ residual correlation. When freely estimated, the value of the residual correlation was $\rho_{i13 \leftrightarrow i19} = 0.455$, with all other estimates similar to those shown in Table 3. Results concerning the factor-based discriminant validity are displayed in Table 4.

Table 4. Differences between factors' square roots of average variances extracted and related factor correlations.

	$\sqrt{AVE_{(f1)}}$ c [a]	$\sqrt{AVE_{(f2)}}$ d [a]	$\sqrt{AVE_{(f3)}}$ e [a]	$\sqrt{AVE_{(f4)}}$ f [a]
	(.774)	(.824)	(.889)	(.755)

$\Phi_{(f1,\ 2)}$ [b]	(.399)	.374 ***	.424 ***		
$\Phi_{(f1,\ 3)}$ [b]	(.401)	.384 ***		.499 ***	
$\Phi_{(f1,\ 4)}$ [b]	(.745)	.025 ***			.006
$\Phi_{(f2,\ 3)}$ [b]	(.381)		.448 ***	.513 ***	
$\Phi_{(f2,\ 4)}$ [b]	(.647)		.177 ***		.107 ***
$\Phi_{(f3,\ 4)}$ [b]	(.481)			.412 ***	.277 ***

* p < 0.05; ** p < 0.01; *** p < 0.001

[a] Square-root of Average Variance Extracted (estimates in brackets); [b] Factor correlations (estimates in brackets). [c] Emotional Neglect. [d] Physical Abuse. [e] Sexual Abuse. [f] Emotional Abuse.

None of the factor correlations exceeded 0.75, which is well below the established 0.85 threshold. Almost all AVE square roots were above the related factor correlations. Formally testing the significance of the difference between the two quantities yielded no significant negative values (entailing $\Phi_{(fj,fk)} > \sqrt{AVE_{(fk)}}$), thus endorsing FDV. Factor-based convergent validity was supported throughout, with all four AVE estimates standing above 0.50 (as implied by squaring $\sqrt{AVE_{(fk)}}$ in Table 4). The corresponding McDonald's ω coefficients were $\omega_{(f1)} = 0.880$, $\omega_{(f2)} = 0.913$, $\omega_{(f3)} = 0.950$ and $\omega_{(f4)} = 0.868$.

3.3. Network Analysis the CTQ-SF

Figure 1 shows the regularized partial correlation network. Consistent with the four-latent factor structure corroborated in the CFA (step 5), the 20 nodes formed a network of four communities, with each item belonging strictly to the theorized cluster (Bernstein et al., 2003). The

strongest edges emerged between the items sharing the same communities. There were three negative associations (represented by the orange edges in the figure) involving items i5 and i20, i23 and i17, and items i27 and i8. Furthermore, the network allows visually assessing eight prominent bridges (as reflected by the thickness of the edges) between items from different communities. These involve: i18 and i19; i8 and i7; i3 and i11; i25 and i15; i25 and i17; i14 and i20; i9 and i27; and i28 and i20.

[Figure 1.]

4. Discussion

The findings from this research showed that Bernstein's (2003) original five-factor model did not hold. Conversely, a 20-item, four-factor solution of the CTQ-SF, excluding all PN items, proved more suitable. This solution showed adequate model fit, highly reliable items, factorial item pertinence and unambiguity, and factor-based convergent and discriminant validities.

Our final model was achieved through a multi-step refinement procedure. Both the initial confirmatory (Step 1) and exploratory (Step 2) analyses suggested that most items of the PN factor exhibited issues, as indicated by weak loadings and high residual variances. In the subsequent step (3), we assessed a model retaining the only PN item with a tenable reliability (i2: "*Got (not) taken care of*"), but the entailing 21-item model carried undesirable properties, including several cross-loadings and high

residual correlations. Thus, the PN-related item was omitted from the system, resulting in a sound final 4-factor solution with 20 items. These findings are consistent with other investigations that have identified similar issues with the psychometric properties of the PN factor among adolescents. Nonetheless, unlike our research, two studies proposed a different solution by reallocating two items of the PN dimension to EN (Aloba et al., 2020; Grassi-Oliveira et al., 2014). Another recent study conducted in China proposed a 4-factor model, removing one item from the PN dimension and merging PN and EN into a single factor (Peng et al., 2023). However, these studies did not thoroughly examine potential factor and item mis-specifications or address the convergent and discriminant validities of their final proposed models.

It is important to note that the issues with the items of the PN factor have been long-standing, as several studies with adult samples have also found similar problems (Bader et al., 2009; Behn et al., 2020; Cheng et al., 2018; Dudek et al., 2015; Karos et al., 2012; Klinizke et al., 2012; Villano et al., 2004; Wingenfeld et al., 2010; Wright et al., 2001). The difficulty in finding acceptable psychometric properties related to the PN factor may stem from the vague definition of negligence and how this construct is captured by the questions (Gerdner & Allgulander, 2009). The CTQ operationalizes PN in terms of whether, during childhood, the respondents had enough to eat (i1), were taken care of by any caregiver (i2), if parents used alcohol or drugs (i4), if they ever wore dirty clothes (i6), and if there was someone to take them to the doctor (i26). At face value, the content validity of these items seems debatable since the

questions may relate to other constructs: items i2 and i6 may relate to social support (Sherbourne & Stewart, 1991); i4 to substance abuse (Smith et al., 2010); i6 to sheer poverty (Gelberg et al., 1995); and i1 to food insecurity (Reichenheim et al., 2016). Although a narrower definition of neglect may hold a lack of parental care and failure to meet children's basic needs (English et al., 2005), the items of the PN dimension could be reflecting a more structural type of violence expressed as social injustice and economic inequalities rather than direct willful or conscious neglect from caregivers (Villano et al., 2004). This theoretical argument gains further traction from empirical evidence showing that endorsement rates of the PN items are higher in low/middle-income than in high-income countries (Stith et al., 2009; Viola et al., 2016). Therefore, caution should be exercised, as these five empirical manifests may measure constructs at a macro level rather than the violence a caregiver perpetrates. Should the broader research community recognize that physical neglect indeed qualifies among the crucial dimensions of child maltreatment, it may be necessary to develop new items that clearly distinguish between caregivers' failures to meet children's physical needs and the damaging conditions of poverty and social exclusion. Alternatively, the dimensional scope could be restricted to only four dimensions, either provisionally while new items are advanced or definitively if PN can be dismissed or a new search effort proves unfeasible.

We also identified a residual correlation between item i13 (*Members of the family (not) looked out for each other*) and i19 (*Family (not) felt*

close) pertaining the EN factor. Noting that both items touch upon family features, this could indicate possible content redundancy at an interpretative level. The result warrants more attention in future studies and a careful resolution if corroborated. Researchers may explore different possibilities, including testing the performance of the most reliable and scalable item of the two; creating a single item by merging the two questions; or developing a new item that unifies their semantic contents.

Lastly, both factor-based convergent (FCV) and discriminant (FDV) validities could be endorsed in our 'final' four-factor model. All square roots of AVE far exceeded the respective factor correlations, underscoring that the retained CTQ-SF factors effectively measured separate aspects of childhood trauma with little overlap. Also, all AVE values were above the recommended threshold of 0.50, indicating that the items are strongly related to their ascribed latent constructs. In line with these findings, all scales showed adequate reliability as conveyed by the McDonald's omega values above 0.70. It is worth noting that, despite previous reports of significant factor correlations (Aloba et al., 2020; Bernstein et al., 2003; Charak et al., 2017), little formal efforts had been devoted to evaluate these essential properties until now.

The network analysis provided innovative insights concerning the item aggregation patterns and the strength of their within and between grouping links. From a substantive stance, the analysis aimed to achieve three primary goals: (1) corroborating the previously uncovered dimensional (factor) structure; (2) illustrating how items are entangled

within their purported factors; and (3) pinpointing the items accounting for the factor correlations identified in the confirmatory factor analysis (regardless of whether FDV has been endorsed). Findings showed that the four emerging clusters were consistent with the ESEM and CFA results from steps 4 and 5. Furthermore, the graphical network further highlighted that EA exhibited a large number of strong bridges with the other aggregates, particularly PA and EN. Notably, three prominent connections involved *calling the child names by family members* (i3), *hitting hard enough to leave bruises* (i11), and *emotional maltreatment* (i25) with both *physical maltreatment* (i15) and *hitting the child badly enough to be noticeable by others* (i17). This finding underscores how emotional and physical aggressions (such as name-calling and hitting) are often part of a broader pattern of abusive behavior within dysfunctional families. The cycle of violence theory posits that different forms of abuse are interrelated and may escalate in severity over time (Fereidooni et al., 2023).

Interestingly, the network also highlighted two negative relationships. One was observed between the EA and SA, expressed by *caregivers showing indifference to the child* (i5), along with less *touching her/him sexually* (i20). The other negative relationship was related to the dimensions of EN and SA, indicated by the increased perception that the *parents wished s/he had never been born* (rejection) (i8) and reduced report of *having been sexually assaulted* (i27). This underscores how caregivers who engage in sexual abuse might exhibit contradictory behaviors, which include providing emotional validation, thus creating a

complex and confusing relationship dynamic for the child. Conversely, caregivers who neglect a child emotionally may not necessarily engage in sexual abuse, possibly because their neglectful behavior is indicative of a general disinterest or absence rather than abusive intent. It is also possible that victims of sexual abuse might interpret certain manipulative behaviors from the caregiver as making them feel important in a twisted form of attention.

Our results should be weighed considering some inherent shortcomings of the CTQ itself. First, the CTQ-SF measures child maltreatment retrospectively, and recalls by participants may have been affected by social desirability and memory. Retrospective self-reports can sometimes lead to underreporting or overreporting of traumatic events due to the passage of time and the influence of current psychological state (Hardt & Rutter, 2004). Additionally, social desirability bias may prompt individuals to minimize or deny experiences of maltreatment, potentially leading to inaccurate information (Baldwin et al., 2019). Regarding the limitations of our study, the sample size restricted the possibility to conduct diagnostic and corroboration analyses in different samples, either through data partitioning or, better, by using data from different studies. Furthermore, our study only included adolescents who were currently enrolled in school. Consequently, we do not know if the performance of the CTQ-SF would remain consistent among adolescents who are no longer in school, such as those who may be working to support their families, those involved in drug trafficking or other illicit activities, or those residing in institutions for vulnerable populations or socio-

educational facilities. We cannot rule out that the physical neglect subscale might perform differently if these groups, typically excluded from school settings. Finally, although the Brazilian version of the CTQ-SF (QUESI) underwent a solid cross-cultural adaptation process (Grassi-Oliveira et al., 2006), it is essential to consider the possibility that linguistic and cultural issues may have influenced the results. Even so, in light of previous findings, we believe that the psychometrics characteristics identified are not peculiar to the QUESI but extend to the inner core of the instrument itself. This calls for replications and, possibly, further developments that go beyond local endeavors.

Despite these limitations, this research provides a comprehensive evaluation of the configural, metric and network structures of the CTQ-SF among Brazilian adolescents. The results support the psychometric properties of a 20-item, four-factor model of the CTQ-SF for retrospectively assessing child maltreatment among adolescents. Future research should replicate such findings with larger and more diverse samples from different linguistic, cultural contexts, and include also adolescents not enrolled in schools. Studies may also evaluate the CTQ-SF's scalability, a property scarcely examined before (Reichenheim & Bastos, 2021).

Reference

Aloba, O., Opakunle, T., & Ogunrinu, O. (2020). Childhood Trauma Questionnaire-Short Form (CTQ-SF): Dimensionality, validity, reliability and gender invariance among Nigerian adolescents. *Child Abuse & Neglect, 101*, 104357.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. Structural Equation Modeling, 16(3), 397-438.

Bader, K., Hänny, C., Schäfer, V., Neuckel, A., & Kuhl, C. (2009). Childhood Trauma Questionnaire – Psychometrische Eigenschaften einer deutschsprachigen Version. *Zeitschrift Für Klinische Psychologie Und Psychotherapie, 38*(4), 223-230.

Baldwin, J. R., Reuben, A., Newbury, J. B., & Danese, A. (2019). Agreement between prospective and retrospective measures of childhood maltreatment: A systematic review and meta-analysis. *JAMA Psychiatry, 76*(6), 584-593.

Behn, A., Vöhringer, P. A., Martínez, P., Domínguez, A. P., González, A., Carrasco, M. I., & Gloger, S. (2020). Validación de la versión en español del Childhood Trauma Questionnaire-Short Form en Chile, en una muestra de pacientes con depresión clínica [Validation of the Childhood Trauma Questionnaire-Short Form in Chile]. *Revista Médica de Chile, 148*(3), 336-343.

Bernstein, D. P., Fink, L., Handelsman, L., Foote, J., Lovejoy, M., Wenzel, K., Sapareto, E., & Ruggiero, J. (1994). Initial reliability and validity

of a new retrospective measure of child abuse and neglect. *American Journal of Psychiatry, 151*(8), 1132-1136.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., Stokes, J., Handelsman, L., Medrano, M., Desmond, D., & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect, 27*(2), 169-190.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: The Guilford Press.

Buss, C., Entringer, S., Moog, N. K., Toepfer, P., Fair, D. A., Simhan, H. N., Heim, C. M., & Wadhwa, P.D. (2017). Intergenerational Transmission of Maternal Childhood Maltreatment Exposure: Implications for Fetal Brain Development. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(5), 373-382.

Charak, R., de Jong, J. T. V. M., Berckmoes, L. H., Ndayisaba, H., & Reis, R. (2017). Assessing the factor structure of the Childhood Trauma Questionnaire, and cumulative effect of abuse and neglect on mental health among adolescents in conflict-affected Burundi. *Child Abuse & Neglect, 72*, 383-392.

Cheng, Y.-C., Chen, C.-H., Chou, K.-R., Kuo, P.-H., & Huang, M.-C. (2018). Reliability and factor structure of the Chinese version of Childhood Trauma Questionnaire-Short Form in patients with substance use disorder. *Taiwanese Journal of Psychiatry (Taipei), 32*(1), 52-62.

Dinno, A. (2009). Implementing Horn's parallel analysis for principal component analysis and factor analysis. *The Stata Journal, 9*, 291-298.

Dovran, A., Winje, D., Overland, S. N., Breivik, K., Arefjord, K., Dalsbø, A. S., Jentoft, M. B., Hansen, A. L., & Waage, L. (2013). Psychometric properties of the Norwegian version of the Childhood Trauma Questionnaire in high-risk groups. *Scandinavian Journal of Psychology, 54*(4), 286-291.

Dudeck, M., Vasic, N., Otte, S., Streb, J., Wingenfeld, K., Grabe, H. J., Freyberger, H. J., Schröder, T., Von Schönfeld, C. E., Driessen, M., Barnow, S., & Spitzer, C. (2015). Factorial validity of the short form of the childhood trauma questionnaire (CTQ-SF) in German psychiatric patients, inmates, and university students. *Psychological Reports, 116*(3), 685-703.

English, D. J., Thompson, R., Graham, J. C., & Briggs, E. C. (2005). Toward a definition of neglect in young children. *Child Maltreatment, 10*(2), 190-206.

Engstrom, M., Liu, G., Santana-Gonzalez, C., Teoh, J. Y., Harms, M., Koy, K., & Quevedo, K. (2021). The impact of child abuse on the neurobiology of self-processing in depressed adolescents. *Neurobiology of Stress, 14*, 100310.

Epskamp, S., Maris, G., Waldorp, L. J., & Borsboom, D. (2018). Network psychometrics. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley*

handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development (pp. 953–986). Wiley Blackwell.

Fereidooni, F., Daniels, J. K., Krause-Utz, A. D., Hagenaars, M. A., Smeets, T., Heins, J., Dorahy, M. J., Emmerik, A. A. P. V., de Jong, P. J., Hoekstra, S., Warrens, M. J., & Lommen, M. J. J. (2023). Childhood maltreatment and adulthood victimization: An evidence-based model. *Journal of Psychiatric Research, 167*, 46-62.

Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 439–492). IAP Information Age Publishing.

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology, 39*(2), 291-314.

García-Fernández, A., Martínez-Cao, C., Sánchez-Fernández-Quejo, A., Bobes-Bascarán, T., Andreo-Jover, J., Ayad-Ahmed, W., Cebriá, A. I., Díaz-Marsá, M., Garrido-Torres, N., Gómez, S., González-Pinto, A., Grande, I., Iglesias, N., March, K. B., Palao, D. J., Pérez-Díez, I., Roberto, N., Ruiz-Veguilla, M., de la Torre-Luque, A., Zorrilla, I., ... García-Portilla, M.P. (2024). Validation of the Spanish Childhood Trauma Questionnaire-Short Form in adolescents with suicide attempts. *Frontiers in Psychology, 15*, 1378486.

Garrusi, B., & Nakhaee, N. (2009). Validity and reliability of a Persian version of the Childhood Trauma Questionnaire. *Psychological Reports, 104*(2), 509-516.

Gelberg, L., Stein, J. A., & Neumann, C. G. (1995). Determinants of undernutrition among homeless adults. *Public Health Reports, 110*(4), 448-454.

Gerdner, A., & Allgulander, C. (2009). Psychometric properties of the Swedish version of the Childhood Trauma Questionnaire-Short Form (CTQ-SF). *Nordic Journal of Psychiatry, 63*(2), 160-170.

Golino, H. F., & Christensen, A. P. (2020). EGAnet: Exploratory Graph Analysis – A framework for estimating the number of dimensions in multivariate data using network psychometrics. R package version 0.9.5. Retrieved April 22, 2021, from https://cran.r-project.org/web/packages/EGAnet/citation.html

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS One, 12*(6), e0174035.

Gorsuch, R. L. (1983). *Factor analysis (2 ed.)*. Hillsdale, NJ: Lawrence Erlbaum.

Grassi-Oliveira, R., Stein, L. M., & Pezzi, J. C. (2006). Tradução e validação de conteúdo da versão em português do Childhood Trauma Questionnaire [Translation and content validation of the Childhood Trauma Questionnaire into Portuguese language]. *Revista de Saúde Pública, 40*(2), 249-255.

Grassi-Oliveira, R., Cogo-Moreira, H., Salum, G. A., Brietzke, E., Viola, T. W., Manfro, G. G., Kristensen, C. H., & Arteche, A. X. (2014). Childhood Trauma Questionnaire (CTQ) in Brazilian samples of different age groups: Findings from confirmatory factor analysis. *PLoS One, 9*(1), e87118.

Hagborg, J. M., Kalin, T., & Gerdner, A. (2022). The Childhood Trauma Questionnaire-Short Form (CTQ-SF) used with adolescents - methodological report from clinical and community samples. *Journal of Child & Adolescent Trauma, 15*(4), 1199-1213.

Hair, J. F., Black, B., Babin, B., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.

Hardt, J., & Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: Review of the evidence. *Journal of Child Psychology and Psychiatry, 45*(2), 260-273.

Hernandez, A., Gallardo-Pujol, D., Pereda, N., Arntz, A., Bernstein, D. P., Gaviria, A. M., Labad, A., Valero, J., & Gutiérrez-Zotes, J. A. (2013). Initial validation of the Spanish childhood trauma questionnaire-short form: Factor structure, reliability and association with parenting. *Journal of Interpersonal Violence, 28*(7), 1498-1518.

Horn, J. L. (1965). A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika, 30*, 179-185.

Innamorati, M., Erbuto, D., Venturini, P., Fagioli, F., Ricci, F., Lester, D., Amore, M., Girardi, P., & Pompili, M. (2016). Factorial validity of the

Childhood Trauma Questionnaire in Italian psychiatric patients. *Psychiatry Research, 245*, 297-302.

Instituto Pereira Passos. (2020). População Residente e Estimada - Brasil, Estado do Rio de Janeiro e Município do Rio de Janeiro e Regiões Administrativas (RA) - 2000/2010/2013-2016/2020 (Vol. 2020). Instituto Pereira Passos.

Karos, K., Niederstrasser, N., Abidi, L., Bernstein, D. P., & Bader, K. (2014). Factor structure, reliability, and known groups validity of the German version of the Childhood Trauma Questionnaire (Short-form) in Swiss patients and nonpatients. *Journal of Child Sexual Abuse, 23*(4), 418-430.

Klinitzke, G., Romppel, M., Häuser, W., Brähler, E., & Glaesmer, H. (2012). Die deutsche Version des Childhood Trauma Questionnaire (CTQ) - psychometrische Eigenschaften in einer bevölkerungsrepräsentativen Stichprobe [The German Version of the Childhood Trauma Questionnaire (CTQ): psychometric characteristics in a representative sample of the general population]. *Psychotherapie, Psychosomatik, medizinische Psychologie, 62*(2), 47-51.

Lauritzen, S. L. (1996). *Graphical models*. New York: Oxford University Press.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Meinck, F., Neelakantan, L., Steele, B., Jochim, J., Davies, L. M., Boyes, M., Barlow, J., & Dunne, M. (2023). Measuring Violence Against Children: A COSMIN Systematic Review of the Psychometric Properties of Child and Adolescent Self-Report Measures. *Trauma, Violence, & Abuse, 24*(3), 1832-1847.

Mizuki, R., & Fujiwara, T. (2021). Validation of the Japanese version of the Childhood Trauma Questionnaire-Short Form (CTQ-J). *Psychological Trauma: Theory, Research, Practice, and Policy, 13*(5), 537-544.

Muthén, L. K., & Muthén, B. O. (1998-2023). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The long-term health consequences of child physical abuse, emotional abuse, and neglect: a systematic review and meta-analysis. *PLOS Medicine, 9*(11), e1001349.

Paquette, D., Laporte, L., Bigras, M., & Zoccolillo, M. (2004). Validation de la version française du CTQ et prévalence de l'histoire de maltraitance [Validation of the French version of the CTQ and prevalence of the history of maltreatment]. *Santé Mentale au Québec, 29*(1), 201-220.

Peng, C., Cheng, J., Rong, F., Wang, Y., & Yu, Y. (2023). Psychometric properties and normative data of the Childhood Trauma Questionnaire-Short Form in Chinese adolescents. *Frontiers in Psychology, 14*, 1130683.

Petrikova, M., Kascakova, N., Furstova, J., Hasto, J., & Tavel, P. (2021). Validation and Adaptation of the Slovak Version of the Childhood Trauma Questionnaire (CTQ). *International Journal of Environmental Research and Public Health, 18*(5), 2440.

R Core Team. (2024). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

RStudio Team. (2024). *RStudio: Integrated Development for R* (Version 2023.12.1 Build 402). RStudio, PBC. Retrieved from https://www.rstudio.com/

Reeve, B. B., Hays, R. D., Bjorner, J. B., Cook, K. F., Crane, P. K., Teresi, J. A., Thissen, D., Revicki, D. A., Weiss, D. J., Hambleton, R. K., Liu, H., Gershon, R., Reise, S. P., Lai, J. S., Cella, D., & PROMIS Cooperative Group. (2007). Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care, 45*(5 Suppl 1), S22–S31.

Reichenheim, M. E., Interlenghi, G. S., Moraes, C. L., Segall-Corrêa, A. M., Pérez-Escamilla, R., & Salles-Costa, R. (2016). A Model-Based Approach to Identify Classes and Respective Cutoffs of the Brazilian Household Food Insecurity Measurement Scale. *Journal of Nutrition, 146*(7), 1356-1364.

Reichenheim, M. E., & Bastos, J. L. (2021). What, what for and how? Developing measurement instruments in epidemiology. *Revista de Saúde Pública, 55*, 40.

Richards, T. N., Tillyer, M. S., & Wright, E. M. (2017). Intimate partner violence and the overlap of perpetration and victimization: Considering the influence of physical, sexual, and emotional abuse in childhood. *Child Abuse & Neglect,* 67, 240-248.

Romano, E., Babchishin, L., Marquis, R., & Fréchette, S. (2015). Childhood Maltreatment and Educational Outcomes. *Trauma, Violence, & Abuse, 16*(4), 418-437.

Sherbourne, C. D., & Stewart, A. L. (1991). The MOS social support survey. *Social Science & Medicine, 32*, 705-714.

Smith, C. A., Elwyn, L. J., Ireland, T. O., & Thornberry, T. P. (2010). Impact of adolescent exposure to intimate partner violence on substance use in early adulthood. *Journal of Studies on Alcohol and Drugs, 71*(2), 219-230.

Sorbom, D. (1989). Model modification. *Psychometrika*, 54, 371–384.

StataCorp. (2019). *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC.

Stith, S. M., Liu, T., Davies, L. C., Boykin, E. L., Alder, M. C., Harris, J. M., & Dees, J. E. M. E. G. (2009). Risk factors in child maltreatment: A meta-analytic review of the literature. *Aggression and Violent Behavior, 14*(1), 13-29.

The Lancet Public Health. (2021). Childhood adversity: a profound determinant of health. *The Lancet Public Health, 6*(11), e780.

Thombs, B. D., Bernstein, D. P., Lobbestael, J., & Arntz, A. (2009). A validation study of the Dutch Childhood Trauma Questionnaire-Short Form: factor structure, reliability, and known-groups validity. *Child Abuse & Neglect, 33*(8), 518-523.

Tussey, B. E., Tyler, K. A., & Simons, L. G. (2021). Poor Parenting, Attachment Style, and Dating Violence Perpetration Among College Students. *Journal of Interpersonal Violence, 36*(5-6), 2097-2116.

United Nations. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. New York: United Nations.

Varlioglu, R., & Hayes, B. E. (2022). Gender differences in the victim-offender overlap for dating violence: The role of early violent socialization. *Child Abuse & Neglect*, 123, 105428.

Villano, C. L., Cleland, C., Rosenblum, A., Fong, C., Nuttbrock, L., Marthol, M., & Wallace, J. (2004). Psychometric utility of the childhood trauma questionnaire with female street-based sex workers. *Journal of Trauma & Dissociation, 5*(3), 33-41.

Viola, T. W., Salum, G. A., Kluwe-Schiavon, B., Sanvicente-Vieira, B., Levandowski, M. L., & Grassi-Oliveira, R. (2016). The influence of geographical and economic factors in estimates of childhood abuse and neglect using the Childhood Trauma Questionnaire: A worldwide meta-regression analysis. *Child Abuse & Neglect, 51*, 1-11.

Waldron, E. M., Howard, K. R., & Reinecke, M. A. (2019). The long-term effect of trauma history on adolescent depression treatment. *Psychological Trauma: Theory, Research, Practice, and Policy, 11*, 751-759.

Wang, X., Ding, F., Cheng, C., He, J., Wang, X., & Yao, S. (2022). Psychometric Properties and Measurement Invariance of the Childhood Trauma Questionnaire (Short Form) Across Genders, Time Points and Presence of Major Depressive Disorder Among Chinese Adolescents. *Frontiers in Psychology, 13*, 816051.

Wingenfeld, K., Spitzer, C., Mensebach, C., Grabe, H. J., Hill, A., Gast, U., ... & Driessen, M. (2010). Die deutsche Version des Childhood Trauma Questionnaire (CTQ): Erste Befunde zu den psychometrischen Kennwerten [The German version of the Childhood Trauma Questionnaire (CTQ): preliminary psychometric properties]. *Psychotherapie, Psychosomatik, medizinische Psychologie, 60*(11), 442-450.

Wright, K. D., Asmundson, G. J., McCreary, D. R., Scher, C., Hami, S., & Stein, M. B. (2001). Factorial validity of the Childhood Trauma Questionnaire in men and women. *Depression and Anxiety, 13*(4), 179-183.