

Psychometric Review of the Childhood Trauma Test

Sabina Alispahić, Ph.D.
Professor of Clinical and Health Psychology, University of Sarajevo

ORCID: 0009-0005-6740-6106 • 29 September 2026

Overall assessment

The IDRIlabs [Childhood Trauma Test](#) shows encouraging internal psychometric properties across 26,934 completed sittings. Its 25 items cover five domains of recalled childhood maltreatment: emotional abuse, physical abuse, sexual abuse, emotional neglect, and physical neglect. Overall consistency is strong, four facets show good to high consistency, and all items relate positively to the scored total.

The findings support cautious optimism about the measure as an educational self-report profile of abuse and neglect experiences. Physical Neglect is the least internally consistent facet and warrants focused review. The evidence does not establish diagnostic accuracy, representative population norms, or independent verification of an individual’s history.

Internal consistency

Reported Cronbach’s alpha is .917 overall. Facet estimates are .85 for Emotional Abuse, .88 for Physical Abuse, .90 for Sexual Abuse, .90 for Emotional Neglect, and .70 for Physical Neglect. Each facet contains five items.

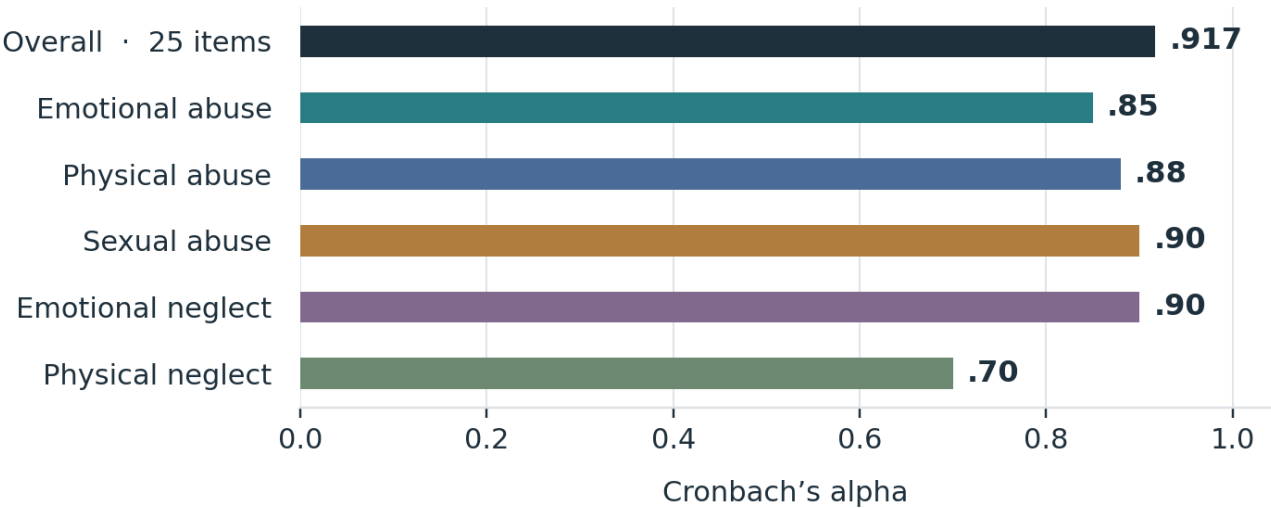


Figure 1. Reported internal consistency for the overall score and five facets.

This pattern supports the coherence of the item set, with greater caution needed for Physical Neglect. Alpha alone does not establish that all items measure a single construct, that the five facets are structurally distinct, or that the scores can classify clinical conditions accurately.

Sample quality and score interpretation

The dataset covers 30 July to 9 September 2026 across 15 published locales. Of 27,039 logged result-page records, 105 same-visitor, identical-answer duplicates were removed, leaving 26,934 gradable sittings. No additional records were excluded as incomplete or out of range.

There are 25,485 sittings with a browser identifier, representing 23,907 distinct identifiers; 1,234 identifiers occur more than once. A further 1,449 sittings lack an identifier. These counts describe website sittings and browser identifiers rather than verified unique people or a representative population sample.

Uniform answers occur in 57 sittings (0.21%) and were retained. Because seven items are reverse-keyed, answering every item identically does not necessarily produce an extreme overall score. The reported uniform-response scores of 28, 50, and 72 agree with the scoring formula. Median completion time is 1 minute 54 seconds in 10,062 timed sittings; no time-based filter was applied.

Score distribution

The overall mean and median are both 42/100, with a standard deviation of 19.4, skewness of 0.18, and excess kurtosis of -0.26 . The middle half of sittings falls between 28 and 55, while the 5th and 95th percentiles are 10 and 75. These values indicate substantial dispersion without a pronounced overall skew.

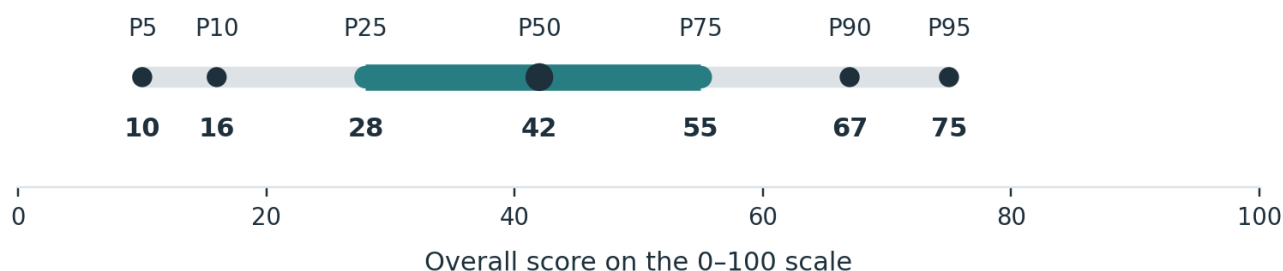


Figure 2. Reported overall-score percentiles. The teal segment marks the middle 50% of sittings.

Scoring consistency

The codebook, reported item means, facet means, and overall mean agree within their rounding precision. No response-range or denominator mismatch is apparent. A percentage score represents a proportion of the possible scored total, not a percentile or probability of maltreatment. The percentiles describe this self-selected website sample and should be treated as sample reference values.

Facet patterns and structural interpretation

Facet means vary appreciably, with the highest average endorsement on Emotional Abuse and the lowest on Sexual Abuse. Differences in content and endorsement patterns mean that these percentages should not be interpreted as interchangeable measures of severity or as population prevalence estimates.

Facet	Items	Mean	SD	Alpha
Emotional Abuse	5	67.3	26.5	.85
Physical Abuse	5	34.0	29.1	.88
Sexual Abuse	5	21.4	28.3	.90
Emotional Neglect	5	53.5	27.0	.90
Physical Neglect	5	33.6	22.4	.70

Table 1. Reported facet statistics on the 0–100 scoring scale.

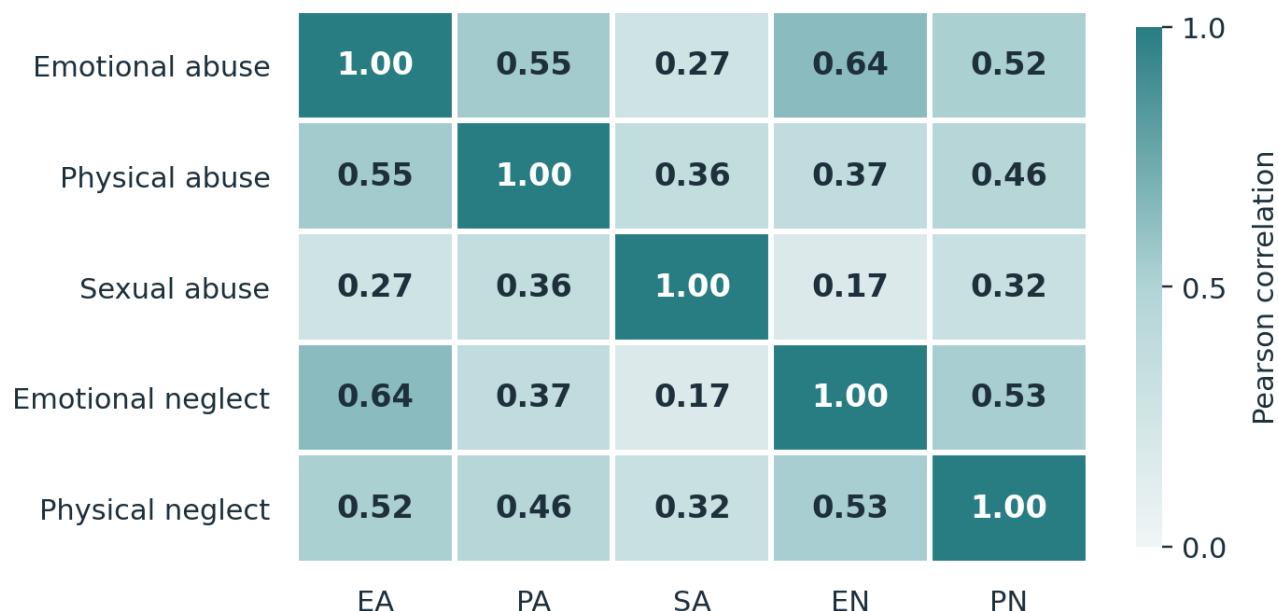


Figure 3. Facet correlations. EA = Emotional Abuse; PA = Physical Abuse; SA = Sexual Abuse; EN = Emotional Neglect; PN = Physical Neglect.

Pairwise correlations range from .17 to .64. Emotional Abuse and Emotional Neglect have the strongest association (.64), while Sexual Abuse and Emotional Neglect have the weakest (.17). The pattern is compatible with related but meaningfully differentiated experiences and supports retaining a facet profile for further evaluation.

Sexual Abuse combines high internal consistency (.90) with relatively modest correlations with the other facets (.17–.36). Its weaker association with the broader total is therefore not evidence of a defective subscale. An overall score can summarize shared variation but should not obscure a high score in an individual domain.

Item performance and content review

Corrected item–total correlations range from .35 to .65, and first-principal-component loadings range from .39 to .72. All values are positive in the scored direction. The seven reverse-keyed items also have positive scored loadings; their negative raw-direction loadings are consistent with the intended reversal.

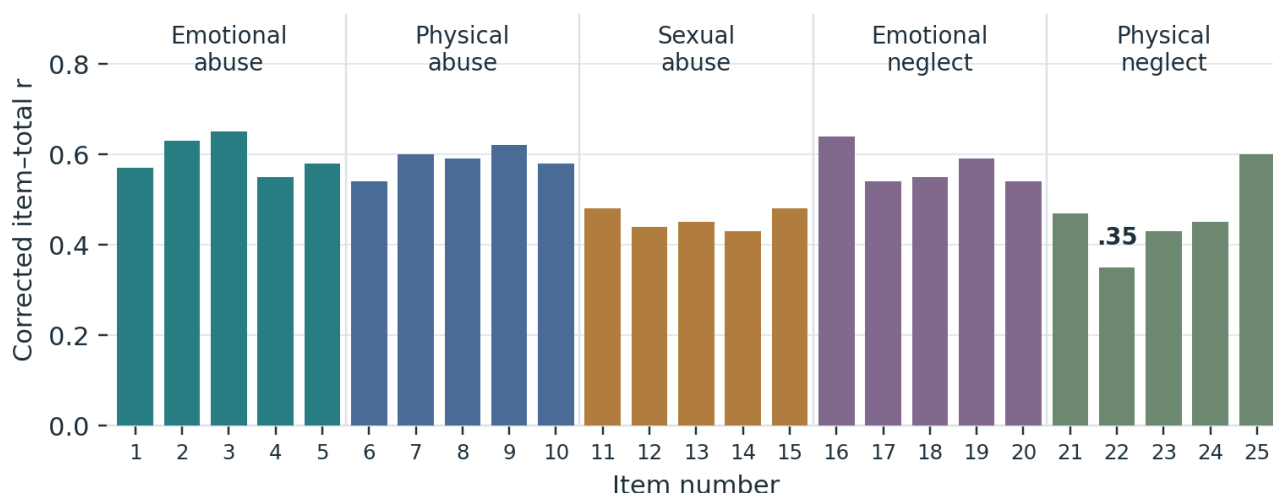


Figure 4. Corrected item–total correlations after reverse-scoring, grouped by assigned facet.

Physical Neglect

Physical Neglect has the lowest facet alpha (.70). Its content spans food, clothing, access to medical care, general care, and parental intoxication. This breadth may contribute to lower internal consistency, although the present summaries cannot determine the reason.

Item 22, concerning parents who were drunk or high, has the lowest overall item–total correlation (.35) and first-component loading (.39). Parental intoxication may indicate a context in which neglect occurs without measuring the same experience as unmet physical needs. Its relevance, wording, and within-facet association should be examined before considering revision. Removing it solely to increase alpha would not establish a better measure.

Interpretation of the item evidence

Sexual Abuse items have corrected item–total correlations of .43–.48 alongside a facet alpha of .90. This combination illustrates why item selection should consider the intended facet as well as the overall total. Positive total correlations are encouraging, but within-facet item discrimination and factor loadings are still needed.

The item content broadly corresponds to the five named domains. The published CTQ-SF contains 25 clinical items plus three validity items; the present item set contains the 25 domain items and does not include that three-item validity component. It should therefore be described as an implementation based on the CTQ-SF rather than assumed to reproduce every feature of the complete instrument.

Conclusion

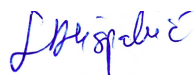
The Childhood Trauma Test shows a promising internal psychometric foundation. Strong overall consistency, good to high reliability in four of the five facets, positive item discrimination, and a broad overall-score distribution support its continued development and cautious educational use. The item and score summaries are arithmetically consistent with the documented response range and reverse-scoring rules.

The main priorities concern the weaker Physical Neglect facet, confirmation of the five-domain structure, and external and cross-language validation. The current evidence supports a coherent profile of self-reported abuse and neglect experiences. It does not establish diagnostic status, maltreatment prevalence, or factual verification of a respondent's history.

Methods and reference

This review evaluates aggregate results for 26,934 sittings, the 25-item codebook, and scoring documentation dated 9 September 2026. Reliability, correlations, percentiles, and item statistics are reported estimates. Checks compared the scoring specification with rounded item and facet means and expected uniform-response scores. Individual response records were not available for independent recalculation, factor analysis, or sensitivity testing.

See also: Bernstein, D. P., et al. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect*, 27(2), 169–190. doi:[10.1016/S0145-2134\(02\)00541-0](https://doi.org/10.1016/S0145-2134(02)00541-0)



Sabina Alispahić, Ph.D.

Date: 30.9.2026.