Psychology International Conference (PSYCIC)

Available online at: **www.sicash.sliit.lk**

Assessment of Judgmental Validity of the Sinhala Physical Abuse Subscale (S-PAS) of the Childhood Trauma Questionnaire–Short Form (CTQ-SF)

Vibhavi Nayanathara[1]*, Naren Deepan Selvaratnam[1, 2]

[1]*School of Psychology, Sri Lanka Institute of Information Technology (SLIIT), Malabe, 10115, Sri Lanka*
[2]*Faculty of Graduate Studies, University of Colombo, Colombo 07, 00700, Sri Lanka*

Corresponding author*: vibhavinayanathara005@gmail.com

Abstract

Childhood physical abuse (CPA) remains a critical public health issue in Sri Lanka, yet there is a notable absence of culturally validated assessment tools in the Sinhala language. This study aimed to translate and assess the content validity of the 5-item Physical Abuse subscale from the Child Trauma Questionnaire–Short Form (CTQ-SF) using the Delphi method. A single round of Delphi was conducted with five subject matter experts (SMEs) selected based on extensive experience in trauma, psychology, and scale validation. Experts rated five Sinhala-translated items against the five content validation criteria of the Delphi method using a 9-point Likert scale. Descriptive statistics, item-level content validity index (I-CVI), scale-level content validity index (S-CVI), and Fleiss' Kappa were used for analysis. All items received high mean ratings ($M > 8.0$) with low standard deviation and coefficient of variation, indicating strong consensus among SMEs. I-CVI values ranged from .80 to 1.00, and S-CVI exceeded .90 across all criteria, demonstrating excellent content validity. One item showed slightly lower agreement among SMEs for the appropriateness of language used, though no items required revision. Fleiss Kappa Statistic further supplemented the SME consensus as the raw agreement for all criteria remained 92% and above. Findings support the cultural and linguistic adequacy of the Sinhala version of the subscale for use in CPA assessment and research.

Keywords: Childhood physical abuse, Delphi method, Content Validity Index (CVI), Fleiss Kappa Statistic (FKS), judgmental validity, psychometrics

Introduction

Childhood physical abuse (CPA) is a significant global public health concern, affecting an estimated 17.3% of children worldwide (Whitten et al., 2023). Recent data from the World Health Organization (2024) indicate a rising prevalence of CPA across diverse regions, underscoring the persistent and widespread nature of this issue. Alarmingly, approximately 60% of children under the age of five are subjected to routine physical punishment and psychological aggression by parents or caregivers. Furthermore, according to a UNICEF (2024) report, nearly 400 million children globally are exposed to physical violence within domestic settings.

Although child physical abuse has garnered significant global attention, the prevalence and impact of this issue in Sri Lanka present distinctive obstacles that necessitate immediate attention. In

https://doi.org/10.54389/WGIB6281

Sri Lanka, child physical abuse persists as a significant issue ingrained in the nation's societal fabric. A study conducted by UNICEF (n.d.) revealed that 40.7% of Sri Lankan parents utilize some form of physical punishment on their children. Furthermore, a UNICEF (n.d.) study indicated that a recent government study among schoolchildren found that 80% of students had experienced at least one form of corporal punishment in schools. A UNICEF (2024) survey indicates that 96% of children experience physical abuse, including actions such as shaking the ears, striking with a cane or belt, and tying or beating them. The primary factors contributing to this physical abuse include cultural norms, economic constraints, and the absence of comprehensive child protection regulations. The National Child Protection Authority of Sri Lanka asserts that corporal punishment is regarded as a conventional and appropriate disciplinary method in numerous cultures (National Child Protection Authority, 2020).

Childhood physical abuse (CPA) continues to be a pervasive concern in Sri Lanka, underpinned by entrenched cultural norms and traditional disciplinary practices. Although global progress has been made in recognizing and responding to CPA, Sri Lanka encounters distinct challenges that impede effective intervention and prevention efforts (Hettiarachchi, 2020). One of the most critical obstacles is the absence of culturally adapted and psychometrically validated assessment tools in the Sinhala language—resources that are essential for accurately identifying and addressing CPA within the Sri Lankan socio-cultural context.

The lack of robust, culturally sensitive assessment instruments significantly hinders the capacity of healthcare professionals, educators, and policymakers to accurately identify, measure, and address cases of childhood physical abuse (CPA). In the absence of tools that are linguistically and culturally appropriate, assessments are prone to inaccuracy or misinterpretation, resulting in underreporting and insufficient support for affected children (Zoysa et al., 2010).

Addressing this gap requires the development and validation of child physical abuse (CPA) assessment tools that are linguistically and culturally appropriate for Sinhala-speaking populations. Such tools would enhance the accuracy of CPA identification, support timely and contextually relevant interventions, and deepen understanding of abuse dynamics in Sri Lanka. Prioritizing culturally responsive instruments is essential for advancing child protection systems and promoting children's well-being. Accordingly, the present study aimed to translate and assess the judgmental validity of the Sinhala version of the 5-item Physical Abuse Subscale of the Child Trauma Questionnaire–Short Form (CTQ-SF) using the Delphi method.

Methodology

Design

This study employed a quantitative, judgment-based validation design using the Delphi method, a technique for achieving consensus among a group of subject matter experts (SMEs) through a structured rating process. The primary objective of using the Delphi technique was to assess the content validity of the items of the Sinhala translation of the Physical Abuse subscale from the Child Trauma Questionnaire–Short Form (CTQ-SF). A single round of SME evaluation was conducted to assess the appropriateness and cultural relevance of the translated items.

Participants

A total of five experts were purposively selected to serve on the content validation panel, based on their extensive knowledge. The inclusion criteria required that each expert possess a minimum of ten years of professional experience in a relevant field, an academic or professional background specifically related to trauma, psychology, or scale validation, and a willingness to participate in the Delphi process. All selected panelists held either a PhD in their relevant fields or demonstrated substantial experience in the assessment and evaluation of child trauma and abuse.

Instrument development

The instrument assessed in the present study was the Sinhala-translated *Physical Abuse* subscale of the *Child Trauma Questionnaire–Short Form (CTQ-SF)*, originally developed by Bernstein et al. (1994) and subsequently refined by Bernstein et al. (2003). This subscale consists of the 9th, 11th, 12th, 15th, and 17th items of CTQ-SF, which operationalize the frequency and severity of experiences related to childhood physical abuse. Scoring is classified into four distinct categories: no abuse (scores ≤7), low (8–9), moderate (10–12), and severe (13–25).

The translation process began with an initial English-to-Sinhala translation conducted by the researcher. This preliminary version was then reviewed and refined by a panel of five subject matter experts (SMEs), who meticulously assessed the linguistic and cultural appropriateness of the items, corrected grammatical inaccuracies, improved sentence structures, and selected terminology that was both culturally and contextually relevant. The finalized Sinhala version of the Childhood Physical Abuse (CPA) subscale was subsequently compared with an independently revised translation to identify any discrepancies in meaning, allowing for further modifications that enhanced the overall translation quality. Following these revisions, the refined Sinhala version underwent content validation through a Delphi procedure involving the same five SMEs.

Delphi procedure

In the single round of the Delphi procedure implemented in this study, SMEs were independently requested to evaluate each of the five items comprising the Sinhala-translated Physical Abuse subscale according to five predetermined content validation criteria: linguistic appropriateness, conceptual assessment, preservation of the original conceptual meaning, suitability for individuals aged 18 years and above, and cultural relevance. Each item was assessed using a 9-point Likert scale, with a score of 1 denoting complete irrelevance and a score of 9 indicating high relevance. Scores ranging from 7 to 9 were interpreted as strong agreement concerning the validity of the item. An item was deemed suitable to be retained if it did not have 70% or more of the expert's ratings in the range of 1-3 in at least one of the five Delphi criteria (De Zoysa et al., 2010).

Data analysis

The collected ratings were compiled and subjected to quantitative analysis to evaluate the extent of SME consensus and to ascertain the appropriateness of each item for inclusion in the final instrument. First, descriptive statistics were calculated for each item across all five criteria. Mean (M) indicated the average level of perceived relevance of items, while the standard deviation (SD) reflected variability among expert responses. The coefficient of variation (CV), obtained by dividing the SD by the mean, was used to standardize this variability. Lower CV values suggested greater consensus among experts, whereas higher values indicated potential disagreement or ambiguity. Next, I-CVI (Item-level Content Validity Index) was calculated for each item across all criteria. I-CVI is the proportion of experts' agreement for a given item within a specific Delphi criterion. An I-CVI of 1.00 denotes complete agreement among all five experts regarding the high relevance of an item, whereas an I-CVI of 0.80 indicates that four out of five experts rated the item as highly relevant. I-CVI further supplements the M, SD, and CV of each item, further indicating the content relevance of the items. Poor I-CVI warrants an item to be subjected to revision and to be assessed by the same set of SMEs in a subsequent round of the Delphi.

Similarly, the average of I-CVI for all 5 items within a specific Delphi criterion was assessed using the content validity index for scale (S-CVI) and S-CVI / (UA) universal agreement. S-CVI offers an aggregated representation of SME consensus. A threshold of .90 or above typically indicates excellent scale-level content validity (Polit & Beck, 2006). S-CVI/(UA), on the other hand, provides a more conservative estimate of content validity by identifying the proportion of items that achieve

unanimous agreement among all SMEs within a particular Delphi criterion. In the present study, greater interpretative weight was assigned to I-CVI and S-CVI due to their stability and reliability in indicating overall content validity.

Table 1: I-CVI, S-CVI, and S-CVI/UA

I-CVI	**S-CVI**	**S-CVI/UA**
$= \frac{\text{Number of YES for an item}}{\text{Total number of criteria}}$	$= \frac{\text{Sum of ICVI}}{\text{Number of items}}$	$\frac{\text{Number of items rated relevant unanimously}}{\text{Total number of items}}$

To further assess the reliability of all expert ratings, the Fleiss Kappa Statistic (FKS) was calculated. FKS can be calculated using the formula displayed in Table 2. Here, **Pe** is the expected agreement while **Po** is the observed agreement between SMEs for all items within a specific Delphi criterion. FKS provides an indication of the overall consistency of SME ratings, which further supplements the overall content validity of the translated items.

Table 2. The breakdown of the Fleiss Kappa Statistic (Landis & Koch, 1977)

K_{FKS} =	**Expected agreement**	**Observed agreement**
$\frac{Po - Pe}{1 - pe}$	Pe = Σp_j^2	Po = $\frac{1}{N \cdot n \cdot (n-1)} (\sum_{i=1}^{N} \sum_{j=1}^{K} n_{ij}^2 - N \cdot n)$

Results and discussion

Following this comparison, a series of minor but meaningful revisions were made to ensure greater conceptual equivalence and linguistic precision. During this process, specific linguistic nuances akin to translation processes were identified and rectified. For instance, terms such as "*so hard*", "*board*", and "*mark*" had multiple possible Sinhala translations, which initially led to differing interpretations. However, following structured discussions and systematic comparisons, the panel reached consensus on the most appropriate and culturally resonant word choices. As a result of this rigorous validation process, all five SMEs assigned high ratings, ranging from 7-9, to each of the five criteria across all items (Table 3), indicating a strong support for content validity of the Sinhala version of the Physical Abuse subscale, hereinafter called Sinhala Physical Abuse Scale (S-PAS).

Table 3. Aggregated expert ratings for Sinhala Physical Abuse Subscale (S-PAS)

	Content-related validation						Consensual-related validation								
CPA Scale Item	Appropriateness of the language used			Assessment of the concept			Retains conceptual meaning			Appropriateness with the individuals of 18 years and above			Cultural relevance		
	1-3	4-6	7+	1-3	4-6	7+	1-3	4-6	7+	1-3	4-6	7+	1-3	4-6	7+
1			100			100			100			100			100
2			100			100			100			100			100
3		20	80			100			100			100			100
4			100			100			100			100			100
5			100			100			100			100			100

Note. All values given above are aggregated percentages

Mean (*M*), standard deviation (*SD*), and coefficient variation (*CV*) for all items across all Delphi criteria were assessed. All items had mean ratings above 8 and *SD* and *CV* below 1, except for items 1, 2, and 3 in the criteria "appropriateness of language used," indicating minor variation in expert consensus for the language usage in those items. Item 3 also had the lowest I-CVI of .80, which is still

acceptable considering the Delphi process included 5 experts. All other items for all categories had I-CVI at 1.0. Similarly, S-CVI for all items was 1.00 across all Delphi criteria except for "appropriateness of language used" which is at .96. S-CVI / UA also displayed unanimous agreement for all items across all criteria in the Delphi process except for "appropriateness of language used" (.80). However, none of the scores indicated any item warranting serious amendments or a requirement for another round of the Delphi process.

Table 4: I-CVI, S-CVI, and S-CVI/UA for Delphi criteria (Round 1 with 5 rates)

	Content-related validation		Consensual-related validation		
	Appropriateness of language used	Assessment of the concept	Retains the conceptual meaning	Appropriateness with individuals of 18 years and above	Cultural relevance
Item 1 I-CVI	1.0	1.0	1.0	1.0	1.0
M ǀ *SD* ǀ*Coef.var*	8.2 ǀ1.1 ǀ 0.1	8.8 ǀ 0.4 ǀ 0.0	8.8 ǀ 0.4 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0
Item 2 I-CVI	1.0	1.0	1.0	1.0	1.0
M ǀ *SD* ǀ*Coef.var*	8.2 ǀ 1.1 ǀ 0.1	8.8 ǀ 0.4 ǀ 0.0	8.8 ǀ 0.4 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0
Item 3 I-CVI	0.8	1.0	1.0	1.0	1.0
M ǀ *SD* ǀ*Coef.var*	8.0 ǀ 1.4 ǀ 0.1	8.6 ǀ 0.9 ǀ 0.1	8.6 ǀ 0.9 ǀ 0.1	9.0ǀ 0.0ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0
Item 4 I-CVI	1.0	1.0	1.0	1.0	1.0
M ǀ *SD* ǀ*Coef.var*	8.6 ǀ 0.9 ǀ 0.1	9.0 ǀ 0.0 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0
Item 5 I-CVI	1.0	1.0	1.0	1.0	1.0
M ǀ *SD* ǀ*Coef.var*	8.4 ǀ 0.9 ǀ 0.1	9.0 ǀ 0.0 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0	9.0 ǀ 0.0 ǀ 0.0
S-CVI	0.96	1.00	1.00	1.00	1.00
S-CVI / UA	0.80	1.00	1.00	1.00	1.00
FKS	-0.042	1.00	1.00	1.00	1.00

***Note*.** Coef. Var - Coefficient of Variation, I-CVI - Content Validity Index for Items, M - Mean, SD - Standard Deviation, S-CVI - Content Validity Index for the Scale, S-CVI/UA - Content Validity Index for the Scale by Universal Agreement, FKS – Fleiss Kappa Statistic

SME ratings for all Delphi criteria demonstrated perfect inter-rater reliability as assessed through the Fleiss Kappa Statistic. However, despite all items receiving I-CVI of 1.0, in the "appropriateness of language used" criteria, the simple drop of I-CVI to .80 in item 3 (indicating that one expert has not agreed about the language usage of the item) yielded a K_{FKS} = -.042. The smaller SME sample and low variability within the criteria (1 disagreement out of a total of 25 decisions), **Pe** (expected agreement) has gotten inflated. What this means for Kappa is that the expert agreement could have been reached simply by guessing, since 24/25 decisions indicate agreement. Here, the total decisions for specific Delphi criteria are derived through $\mathbf{Pe} = \boldsymbol{\Sigma p_j^2}$ where 5 experts per item times the number of items has yielded the sum of 25 decisions. In the present study, **Pe** for "appropriateness of language used" criteria is .9232, and the **Po** (observed agreement) is .92. Hence, K_{FKS} = Po – Pe / 1 – Pe has generated a negative value. Here, the negative value does not reflect a genuine inter-rater discord, but rather an artifact of mathematical calculation, not of expert behavior. However, Po = .92 still implies that the raw agreement for the items in the scale by experts is highly acceptable (92%).

Conclusion

This study demonstrates that the Sinhala-translated Physical Abuse subscale of the CTQ-SF satisfies all essential psychometric criteria for content validity, as evidenced by robust I-CVI and S-CVI values alongside strong consensus among subject matter experts (SMEs). These findings affirm the scale's appropriateness for assessing childhood physical abuse (CPA) within the Sri Lankan cultural and linguistic context. The establishment of a validated instrument has significant implications for clinical

practice and psychological assessment, facilitating precise identification and evaluation of CPA cases. Furthermore, it supports the integration of culturally sensitive measures in trauma assessment, ultimately enhancing the accuracy of diagnoses and informing targeted intervention strategies. Adoption of this tool can contribute to advancing evidence-based clinical care and research focused on childhood trauma in Sri Lanka.

Reference

Bernstein, D. P., Fink, L., Handelsman, L., & Foote, J. (1994). Childhood Trauma Questionnaire. *PsycTESTS Dataset*. https://doi.org/10.1037/t02080-000

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. https://doi.org/10.1016/s0145-2134(02)00541-0

De Zoysa, P., Rajapakse, L., & Newcombe, P. A. (2010). Adaptation and validation of the Parent-Child Conflict Tactics Scale for use in Sri Lanka. *Ceylon Medical Journal*, *50*(1), 11. https://doi.org/10.4038/cmj.v50i1.1582

Hettiarachchi, D. H. (2020). Management of victims of child abuse in Sri Lanka: The view of a Child Psychiatrist. *Sri Lanka Journal of Child Health*, *49*(3), 279. https://doi.org/10.4038/sljch.v49i3.9148

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

National Child Protection Authority. (2024). *Home | National Child Protection Authority*. Childprotection.gov.lk. https://childprotection.gov.lk/en/

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

UNICEF. (n.d.). *Child protection in middle childhood*. Www.unicef.org. https://www.unicef.org/srilanka/child-protection-middle-childhood

UNICEF. (2024). *Nearly 400 million young children worldwide regularly experience violent discipline at home – UNICEF*. Www.unicef.org. https://www.unicef.org/press-releases/nearly-400-million-young-children-worldwide-regularly-experience-violent-discipline

Whitten, T., Tzoumakis, S., Green, M., & Dean, K. (2023). Global Prevalence of Childhood Exposure to Physical Violence within Domestic and Family Relationships in the General Population: A Systematic Review and Proportional Meta-Analysis. *Trauma, Violence, & Abuse*, *25*(2), 152483802311791-152483802311791. https://doi.org/10.1177/15248380231179133

World Health Organization. (2024, November 5). *Child Maltreatment*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/child-maltreatment