**Journal of Multidisciplinary and Translational Research (JMTR)**

journal homepage: https://journals.kln.ac.lk/jmtr/

# Refinement of Scale Items for a Screening Tool for Specific Learning Disorders in Sinhala-Medium Primary School Children: A Delphi Study

W.M.L.D.J. Wijesekara[1]*, A.P. Hewamalage[1], P.K.S. Godamunne[2], U.G. Karunaratne[1], and N.H. Nanayakkara[1]

[1]*Family Health Bureau, Ministry of Health, Sri Lanka,*
[2]*Department of Medical Education, University of Kelaniya, Sri Lanka*

## Abstract

Dyscalculia, Dyslexia, and Dysgraphia are under the umbrella term Specific Learning Disorders (SLDs), which afflict 5 to 15 % of school-aged children globally. Children with SLDs experience mental health issues, social challenges, and academic underachievement. Despite the increasing awareness, Sri Lanka does not have a screening instrument that is culturally and linguistically appropriate for early identification of SLDs. This Delphi survey aimed to achieve non-statistical item reduction through consensus on the relevance, clarity, and contextual suitability of the screening tool items for primary school children. A 50-item pool was initially developed through desk reviews, in-depth interviews, and focus group discussions. Subsequently, a Delphi survey was conducted with a multidisciplinary panel of 15 national and international experts to improve the validity and application. In the initial round, test items were refined by quantitative analysis utilizing the Validation Ratio Grid (VRD) and the Content Validity Ratio (CVR) in addition to qualitative interviews allowing retaining, revising, or eliminating test items. The second round consisted of expert reviews on the refinements, leading to a final 39-item pool assessing abilities in arithmetic (11 items), reading (17 items), and writing (11 items). The Findings underline the need for expert consensus in establishing culturally sensitive assessments. By aligning the tool with the educational context of Sri Lanka, this study advances special education by providing a validated, reliable screening tool for early SLD detection in primary school children. Future research is focused on pilot testing, field validation, and large-scale implementation to assess the effectiveness of this tool.

**Keywords:** Delphi Study, Dyscalculia, Dysgraphia, Dyslexia, Screening Tool

Article info

Article history:
Received 3rd February 2026
Received in revised form 20th March 2026
Accepted 15th April 2026
Available online 27th April 2026

ISSN (E-Copy): ISSN 3051-5262
ISSN (Hard copy): ISSN 3051-5602
Doi:
ORCID iD:
*Corresponding author:
E-mail address: drlasanthawijesekara@gmail.com
(W.M.L.D.J. Wijesekara)

# Introduction

Persistent barriers to reading, writing, and calculation which cannot be attributed to sensory impairments, intellectual disability, or inadequate educational exposure are the hallmarks of SLDs (American Psychiatric Association, 2013). According to Fletcher et al. (2019), inadequacies in executive functioning, working memory, and phonological processing underlie academic underachievement and failure across broader domains of life, such as social and vocational functioning. Long-term consequences of these challenges include school dropouts, limited career opportunities, and reduced quality of life, underscoring the importance of early detection and support (Aro et al., 2018).

SLDs are common among children (Scaria et al., 2022), and according to statistics, 5% to 15% of children in schools worldwide are affected, though this varies depending on culture and linguistic background (Snowling et al., 2011). Data on SLD prevalence in Sri Lanka are limited, indicating a wide knowledge and research deficit (Menikdiwela et al., 2021). Early identification is crucial, as timely interventions have been shown to ease adverse outcomes. Evidence-based programs such as phonological training for Dyslexia, number sense interventions for Dyscalculia, and other school-based interventions have demonstrated efficacy in enhancing learning outcomes when implemented during the early school years (Grimes and Cruz, 2021).

Despite these advancements, Sri Lanka faces unique challenges in addressing SLDs. Limited awareness among school teachers and parents, coupled with a lack of culturally relevant assessment tools, hinders early identification and intervention. Inclusive education policies advocate for the integration of children with SLDs into mainstream classrooms, yet the success of such frameworks depends on the availability of reliable screening instruments and teacher training (Kochhar and Heishman, 2010). The primary school represents the earliest and most appropriate stage for identifying SLDs, as foundational literacy and numeracy skills are typically acquired during this period (Bozatli et al., 2024). Failure to address learning difficulties at this stage can intensify educational inequities and social exclusion (American Psychiatric Association, 2013). Conversely, the Sinhala language, which is a unique Indo-Aryan language with a distinct phonetic system, complex agglutinative morphology, and a syllabic script, is structurally different from many widely used languages worldwide. As such, existing instruments, prominently designed in Western contexts, may not align with the linguistic and educational distinctions of children using the Sinhala medium (Garland, 2005; Peña, 2007). Therefore, the development of culturally and linguistically appropriate screening tools is a pressing need.

This survey aims to develop a screening tool for SLDs in primary school children through a comprehensive process. The methodology includes a desk review, in-depth interviews, and focus group discussions to identify constructs (Arithmetic, Reading, and Writing) and their representative domains and to generate the initial item pool. The pool is then refined through a Delphi survey that gathers expert consensus. Subsequently, a pilot study to test the usability of the instrument and clarity, followed by a community-level survey for factor analysis and validation, was employed (Redding et al., 2006). This script focuses on the Delphi survey stage of the research process.

Delphi surveys are known to be an effective technique for achieving consensus in disciplines where expert judgment is essential or where empirical support is lacking. Humphrey-Murto et al. (2017) highlight the flexibility and structured approach of the Delphi method, which fosters iterative feedback and collective expertise. Similarly, Hasson et al. (2000) emphasize its reliability in generating consensus through repeated rounds of expert input, making it suitable for instrument development and refinement. Furthermore, Powell (2003) emphasises the applicability of the Delphi method in educational and healthcare research, mainly when diverse stakeholder perspectives are vital to the process of refining complex instruments and frameworks improving the reliability and applicability of the final product.

The Delphi process in the present study involved four key steps: (1) Identification of the Problem/Solution (identification of items for the constructs), (2) Recognition of the Expert Panel, (3) Performing Delphi Rounds, and (4) Reaching Consensus (Humphrey-Murto et al., 2017).

Objectives of the Delphi in this survey are to ensure item validity (whether the item is suitable to assess specific skills including arithmetic, reading, and writing), cultural validity (whether the item is relevant for children in Sri Lanka, and sample validity (whether the item is suitable for children between 5 and 8).

# Methodology

## Development of the Initial Screening Items

Before starting the Delphi consensus process in this multi-stage survey, we followed several theory-based steps to develop the initial set of screening tool items. First, we did a thorough desk review using international classification systems and research on Specific Learning Disorders (SLDs) like dyslexia, dysgraphia, and dyscalculia, guided by the DSM-5 and ICD-11 frameworks (American Psychiatric Association, 2013; World Health Organization, 2019). This helped us find key concepts, cognitive areas, and classroom signs linked to early literacy and numeracy problems. Next, we reviewed existing screening and assessment tools, such as the WISC, WIAT, DIBELS, and early-grade curriculum-based measures, to identify suitable task types and scoring methods. We then turned the first set of items into tasks for teachers to use and refined them through interviews and group discussions with psychologists, speech therapists, educators, doctors, and policymakers. Boateng et al. (2018) noted that expert review and repeated feedback are important for making sure new screening tools are valid, fit the context, and work well in practice. Together, these steps ensured the items used in the Delphi phase were grounded in theory, relevant to the context, and practical for classroom screening.

## Participant Selection and the Ethical Clearance Procedures

The study was conducted with who are competent in Sinhala language skills. The experts were selected based on their specialized knowledge of SLDs and experience with the target population according to a specified inclusion and exclusion criteria.

The inclusion criteria required individuals to possess professional expertise in fields relevant to SLDs and related neurodevelopmental disorders. In addition, participants were required to have

practical experience in working with primary school children or in the assessment and management of learning difficulties. The exclusion criteria included individuals with limited professional experience in the relevant domains or those not directly involved in child assessment, intervention, or educational practice. Experts who were unable to commit and did not respond were also excluded. This ensured that the expert panel consisted of individuals with sufficient knowledge and experience to provide informed judgments during the Delphi process.

The experts' willingness to participate was requested in a detailed email. The experts who responded that they could not attend or did not respond to the email were excluded and replaced by other experts. Finally, a total of 15 experts participated in the Delphi study, including a linguist, two consultant child psychiatrists, a child psychologist, an educational psychologist, a speech and language therapist, a consultant community physician, four special education teachers, and four primary school teachers. Data collection was conducted through individual interviews, either face-to-face or via video conferencing, followed by email communication.

During this data collection process, no personal information was collected. As such, at this stage of the study, formal ethical clearance was not required. Furthermore, the study involved a Delphi survey in which subject-matter experts provided professional judgments rather than personal or sensitive data. Participation was entirely voluntary, and all experts were fully informed about the study objectives, procedures, and their role prior to participation. Written informed consent was obtained from all participants. No identifiable personal information was collected, and responses were analysed in aggregate form to ensure anonymity. In line with standard research ethics practices, studies involving expert consultation with minimal risk and no collection of personal data are generally considered exempt from formal ethical review. However, subsequent pilot testing and field testing were done after obtaining ethical clearance from the University of Kelaniya.

## Study Design

The study employed a modified two-round Delphi technique involving the above-mentioned panel of 15 experts in psychology, education, and child development. In the first round, the experts independently evaluated the relevance, clarity, and cultural appropriateness of each item using a structured rating scale. Quantitative indices, including the Content Validity Ratio (CVR) and Content Validity Index (CVI), were calculated to determine item retention. In the second round, the items that did not reach predefined consensus thresholds were revised and resubmitted for re-evaluation.

The CVR and the CVI are widely recommended quantitative approaches in instrument development. The CVR, originally proposed by Lawshe (1975), assesses the extent to which experts judge an item as essential. Experts rated each item on a three-point scale (essential, useful but not essential, not necessary). The obtained CVR values were compared against minimum acceptable thresholds based on panel size (minimum acceptable CVR = 0.49 for 15 experts), as recommended by Ayre and Scally (2014). Items failing to meet the required CVR threshold were flagged for revision or removal. The CVI was used to assess item-level relevance and clarity. Experts rated each item on a Likert scale, and the Item-Level CVI (I-CVI) was calculated as the proportion of experts who rated the item as relevant. A minimum I-CVI value of 0.80 was

considered acceptable, in line with established guidelines (Polit & Beck, 2006; Polit & Beck, 2007). The combined use of CVR and CVI provided a comprehensive assessment of both the essentiality and relevance of items, thereby strengthening the content validity of the screening tool.

## Study Instruments

The study instruments comprised the Delphi questionnaire to assess key aspects, such as face validity, ease of administration, scoring complexity, time consumption, alignment with the curriculum's learning objectives, and construct validity, which was assessed using a Content (Table 1). In addition, it contained a standard form to collect opinions on the Content validity of each item. The Validation Ratio Grid (VRG) was used to assess content validity. A sample form for items in Writing is shown (Figure 1). Similar forms were used for Arithmetic and Reading separately.

**Table 1:** Key aspects of the Delphi questionnaire

| Key Aspect | Question/ Description |
|---|---|
| Face Validity | Does the item effectively measure the intended skill for the age group in the Sri Lankan context? |
| Administration Ease | How easily primary school children can respond, and teachers can administer this item? |
| Complexity of Scoring | How easily a primary class teacher could interpret and assign scores for this item? |
| Time Consumption | What is your idea about the time consumed by this item? |
| Alignment of Learning Objectives | Does this item align with the learning objectives set by teachers' guide? |
| Construct Validity | Is this item essential, somewhat essential, or not essential in the tool? |
| Additional comments or Suggestions | Provide feedback on the development and implementation of the screening tool in schools. |

| Question Code | Essential | Somewhat essential | Not Essential |
|---|---|---|---|
| WI-1 | | | |
| WI-2 | | | |
| WI-3 | | | |
| WI-4 | | | |
| WG-1 | | | |
| WG-2 | | | |
| WG-3 | | | |
| WG-4 | | | |

**Figure 1.** Sample form (writing) used for the collection of quantitative data in round 1 on content validity.

## Data Collection

The entire Delphi process, from sending invitations to the completion of data analysis, was conducted from November 2023 to July 2024.

The Delphi process was conducted in two rounds. In the first round 15 experts were involved and Qualitative and Quantitative feedback regarding the initial set of test items was gathered. In order to ensure the possibility of receiving unbiased and original inputs, the interviews were conducted individually. Each expert was provided with the printed test items and the marking scheme to

review while commenting on each test item. Finally, opinions were gathered about the overall pool of items. The principal investigator conducted the interviews with the Delphi questionnaire, which guided the discussions on the objectives of the study. Parallelly, research assistants were assigned to document all responses, observations, and suggestions provided during the interviews. The inputs were noted down on a printed draft test paper during the interview.

The experts were encouraged to provide detailed feedback for each item on the relevance, clarity, and cultural appropriateness of the test items in the first round of the Delphi Survey. This round was qualitative with mostly open-ended discussions. The items' linguistic and contextual suitability for Sinhala medium primary school students was specifically considered. In addition, experts were invited to point out any possible uncertainties or gaps in the test items and recommend any necessary changes or more material.

Then towards the latter part of round One quantitative feedback was gathered for content validity analysis to complement the qualitative feedback (Figure 1), All the experts categorized each item as "essential," "useful but not essential" (somewhat essential), or "not essential" for inclusion in the screening tool.

Round Two of the Delphi process was conducted following a thorough analysis and revision of the feedback gathered in Round One. The revised set of test items was shared with the expert panel via email, ensuring a convenient and efficient mechanism for review. The experts were asked to review the modified test items critically and provide their input on whether the revisions adequately addressed their earlier concerns. Twelve out of the 15 experts have appropriately replied to the second round. Along with the revised items, a detailed summary of the changes made and the rationale behind each modification was provided.

## Data Analysis

Based on the expert feedback, the research team engaged in a systematic process of retaining, revising, or eliminating test items. Items that met the predefined criteria for validity and cultural relevance were retained, while those considered uncertain or irrelevant were either revised or excluded.

Data analysis involved 1. Refining the items based on qualitative feedback gathered during the rounds, and 2. Applying the Validation Ratio Grid. Validation Ratio was used to quantify expert consensus by calculating the CVR for each test item.

## Steps to Use the Validation Ratio Grid

1. Develop the Item Pool
2. Collect Expert Ratings using a Likert scale (e.g., "essential" vs. "not essential") for each item.
3. Calculate the CVR for each item,
4. Use the Validation Grid,
5. Iterate and finalize (Khodyakov et al., 2023).

The CVR values were calculated using the formula, **CVR = (2ne / N) – 1**, where 'ne' represents the number of experts who rated the item as "essential", whereas the number of experts who participated in the Delphi procedure is denoted by 'N'.

A higher degree of agreement among experts evaluating the items' essentiality is shown by positive CVR values, which range from -1 to +1. A CVR around zero or negative indicates a lack of consensus or agreement that the item is not vital. For example, a CVR of +1 indicates widespread agreement that the item is essential (Lim & Park, 2025; Romero et al., 2023). Accordingly, items with high CVR values were considered 'essential' and 'retained' without further modifications, whereas items with moderate CVR values indicated partial agreement, necessitating improvement. Conversely, items with low CVR values were considered 'non-essential' or 'irrelevant' and were subsequently 'eliminated'.

# Results

The original 50 - item pool was reduced to 39 items while maintaining comprehensive coverage of domains of arithmetic, reading, and writing skills (Tables 2, 3, and 4).

## Item reduction in Arithmetic

**Table 2:** Analysis of arithmetic questions for item reduction

| Item | Experts | | | | | | | | | | | | | | | CVR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | e1 | e2 | e3 | e4 | e5 | e6 | e7 | e8 | e9 | e10 | e11 | e12 | e13 | e14 | e15 | |
| AI-1 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| AI-2 | NE | E | SE | E | E | E | E | E | NE | E | E | SE | E | E | E | 0.47 |
| AI-3 | E | E | E | E | E | E | E | E | NE | E | E | E | E | E | E | 0.87 |
| AG-1 | E | E | E | E | E | E | E | E | E | SE | E | E | E | E | E | 0.87 |
| AG-2 | E | E | E | E | E | E | E | E | E | E | NE | E | E | E | E | 0.87 |
| AG-3 | E | E | E | E | E | SE | E | E | E | E | E | E | E | E | E | 0.87 |
| AG-4 | E | NE | NE | E | NE | E | SE | E | E | SE | E | E | E | E | E | 0.33 |
| AG-5 | E | E | SE | E | E | E | E | E | E | E | NE | E | E | E | E | 0.87 |
| AG-6 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 0.87 |
| AG-7 | E | E | E | E | E | E | E | NE | E | E | E | NE | E | E | E | 0.73 |
| AG-8 | E | E | E | E | NE | E | E | E | E | E | NE | E | E | E | E | 0.73 |
| AG-9 | E | E | E | E | E | E | E | E | NE | E | NE | NE | E | E | E | 0.6 |
| AG-10 | E | E | E | E | E | E | E | E | NE | E | E | E | SE | NE | E | 0.6 |
| AG-11 | E | E | E | E | E | E | E | E | E | E | E | E | NE | E | E | 0.87 |
| AG-12 | E | NE | E | E | SE | E | E | E | E | E | E | E | NE | E | E | 0.73 |
| AG-13 | E | NE | E | E | E | E | E | NE | E | E | NE | NE | E | E | NE | 0.33 |
| AG-14 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| AG-15 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| AG-16 | E | E | E | E | NE | E | E | E | E | E | E | E | E | E | E | 0.87 |

***Note:*** AI – Arithmetic items for Individual Administration (AI-1 to AI-3),
AG – Arithmetic items for Group Administration (AG-1 to AG-16)
e1 to e15 in the first row the 15 experts
E = Essential SE = Somewhat essential NE = Not essential

The initial pool of arithmetic items included 19 questions (**Annexure 1 and Annexure 2).** Three were designated for individual administration (AI-1 to AI-3) and sixteen for group administration (AG-1 to AG-16). Each item was assessed by 15 experts who rated it as E – Essential, SE – Somewhat Essential or NE – Not Essential. The table was used to calculate the content validity ratio (CVR) of each item to determine consensus (Table 2).

Three items (AI-1, AG-14, and AG-15) achieved unanimous expert agreement as Essential, resulting in a perfect CVR of 1.00. Eight items (AI-3, AG-1, AG-2, AG-3, AG-5, AG-6, AG-11 and AG-16) also received a high CVR of 0.87, demonstrating strong agreement. All 11 items were retained, and some were modified based on expert opinions.

Items AG-7, AG-8 and AG-12 (CVR = 0.73), AG-9 and AG-10 (CVR = 0.6), AG-4 and AG-13 (CVR = 0.33) and AI-2 (CVR = 0.47) were reviewed closely and removed due to issues in clarity, redundancy or misalignment with the curriculum (**Annexure 3 and Annexure 4)**.

**Table 3:** Analysis of reading questions for item reduction

| Items | Experts | | | | | | | | | | | | | | | CVR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | e1 | e2 | e3 | e4 | e5 | e6 | e7 | e8 | e9 | e10 | e11 | e12 | e13 | e14 | e15 | |
| RI-1 | E | E | E | E | E | E | E | E | E | E | E | E | E | NE | E | 0.87 |
| RI-2 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| RI-3 | NE | NE | E | E | E | E | E | E | NE | NE | NE | NE | NE | E | NE | -0.067 |
| RI-4 | SE | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 0.87 |
| RI-5 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| RI-6 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| RI-7 | E | E | E | E | E | E | E | E | NE | NE | E | E | E | E | E | 0.73 |
| RG-1 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| RG-2 | E | E | E | E | NE | E | NE | E | E | E | E | SE | E | E | E | 0.6 |
| RG-3 | E | E | E | E | NE | E | E | E | NE | NE | E | E | E | E | E | 0.47 |
| RG-4 | E | E | E | E | E | E | SE | E | E | E | E | E | E | E | E | 0.87 |
| RG-5 | E | E | E | SE | NE | E | E | E | NE | E | E | E | E | E | E | 0.73 |
| RG-6 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| RG-7 | E | E | E | E | E | E | E | NE | E | E | NE | SE | E | E | E | 0.6 |
| RG-8 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| RG-9 | E | E | E | E | E | E | E | NE | E | E | E | E | E | E | E | 1 |
| RG-10 | E | NE | E | E | E | E | E | E | E | E | NE | E | E | E | E | 0.6 |

***Note:*** RI – Reading items for Individual Administration (RI-1 to RI-7)
RG – Reading items for Group Administration (RG-1 to RG-10)
e1 to e15 in the first row the 15 experts
E = Essential SE = Somewhat essential NE = Not essential

## Item reduction in Reading

The initial pool contained seven items for individual administration (RI-1 to RI-7) and 10 for group administration (RG-1 to RG-10) (**Annexure 5 and Annexure 6),** which were rated by the experts, and the CVR was calculated (Table 3).

Seven items (RI-2, RI-5, RI-6, RG-1, RG-6, RG-8 and RG-9) achieved a perfect CVR of 1.00. Three other items (RI-1, RI-4 and RG-4) received a high CVR of 0.87, indicating strong expert consensus. Other items got CVR values as follows: RI-3 (-0.067), RI-7 (0.73), RG-2 (0.6), RG-3 (0.47), RG-7 (0.6), and RG-10 (0.6).

All 7 items that obtained a CVR of 1.00 were retained. From the items obtained, CVR = 0.87, RI-1 and RG-4 were retained while RI-4 was modified. Other items were reviewed based on CVR values and qualitative inputs of the experts and assessed for concerns regarding ambiguity, redundancy or misalignment with expected literacy components for the primary school age group in concern. This analysis resulted in the removal of RI-3, RI-7, and RG-7, and the modification of RG-2, RG-3, and RG-10. Three more items were added. As such, the number of items remained the same after the refinement (17 items) (**Annexure 7 and Annexure 8)**.

## Item Reduction in Writing

The initial pool for Writing had 14 items. 4 items for individual administration (WI-1 to WI-4) and 10 for group administration (WG-1 to WG-10) (**Annexure 9 and Annexure 10)**, which were rated by the experts, and the CVR was calculated.

Three items (WI-4, WG-1, and WG-10) achieved unanimous consensus as essential for a CVR of 1.00. Two other items (WI-3 and WG-3) received a high CVR of 0.87, indicating strong expert consensus. The other items had lower CVR values (Table 4).

All the items were reviewed based on the CVR value and qualitative inputs of the experts. Ambiguity, redundancy or misalignment with expected literacy components for the primary school age group were examined. Two items with a CVR of 1.00 (WI-4 and WG-10) were retained, and WG-1 was excluded as WI-1 was of similar scope. Among the items with a CVR of 0.87, WG-3 was retained, while WI-3 was excluded due to its poor alignment with the curriculum (too easy). Similarly, WG-2 (CVR -0.067) was excluded because it was too easy for the age category. Modification of items was done by combining WG-5 (CVR = 0.2) with WG-6 (CVR = 0.73) and WG-7 (CVR = 0.6) with WG-8 (CVR = 0.73). The final number of items in this pool was 11 (**Annexure 11 and Annexure 12)**.

**Table 4:** Analysis of writing questions for item reduction

| Items | Experts | | | | | | | | | | | | | | | CVR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | e1 | e2 | e3 | e4 | e5 | e6 | e7 | e8 | e9 | e10 | e11 | e12 | e13 | e14 | e15 | |
| WI-1 | SE | E | SE | E | E | E | E | NE | E | E | E | SE | E | NE | E | 0.33 |
| WI-2 | E | E | E | E | E | SE | E | E | NE | E | NE | E | E | E | E | 0.6 |
| WI-3 | E | E | E | E | E | E | E | NE | E | E | E | E | E | E | E | 0.87 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WI-4 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| WG-1 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |
| WG-2 | NE | NE | SE | E | NE | E | E | SE | E | E | E | NE | NE | E | NE | -0.067 |
| WG-3 | E | E | E | E | E | E | E | E | E | E | E | E | NE | E | E | 0.87 |
| WG-4 | NE | SE | E | E | NE | E | E | E | E | E | E | E | E | E | E | 0.6 |
| WG-5 | NE | NE | E | E | NE | E | E | E | NE | E | SE | SE | E | E | E | 0.2 |
| WG-6 | E | E | E | E | NE | E | E | E | NE | E | E | E | E | E | E | 0.73 |
| WG-7 | NE | E | E | E | E | E | E | NE | E | NE | E | E | E | E | E | 0.6 |
| WG-8 | E | E | E | E | E | E | E | E | E | E | NE | NE | E | E | E | 0.73 |
| WG-9 | E | E | E | E | E | E | E | NE | E | E | NE | E | E | E | E | 0.73 |
| WG-10 | E | E | E | E | E | E | E | E | E | E | E | E | E | E | E | 1 |

***Note:*** WI – Writing items for Individual Administration (WI-1 to WI-4)
WG – Writing items for Group Administration (WG-1 to WG-10)
e1 to e15 in the first row the 15 experts
E = Essential SE = Somewhat essential NE = Not essential

Disagreements among experts during the Delphi process were systematically examined using both quantitative and qualitative approaches. In the first round, items that did not meet the predefined consensus thresholds—based on CVR and I-CVI—were identified for further review. The distribution of expert ratings was analyzed alongside qualitative comments to understand the sources of disagreement, which commonly related to item clarity, contextual relevance, or developmental appropriateness.

Items with strong conceptual relevance but lower agreement were revised to address ambiguities in wording or to better align with the Sri Lankan primary school curriculum. These revised items were then re-presented to the expert panel in the second Delphi round for re-evaluation. In contrast, items that consistently demonstrated low CVR values and limited theoretical relevance were removed from further consideration.

Final item selection was based on achieving the predefined consensus criteria (CVR above the minimum acceptable threshold for panel size and I-CVI ≥ .80), in conjunction with qualitative agreement among experts regarding the item's clarity and contextual suitability.

With all the refinements, the final pool of items was 39 in total, adequately covering the areas required to screen SLDs in primary school children (Table 5).

**Table 5:** Item reduction by Delphi survey

| Worksheet / Test | Version 1 No. of Items (Before Delphi) | Version 2 No. of Items (After Delphi) |
|---|---|---|
| Arithmetic | 19 | 11 |
| Reading | 17 | 17 |
| Writing | 14 | 11 |
| **Total** | **50** | **39** |

## Discussion

Cultural and linguistic relevance is important in adapting assessment tools to ensure their validity and reliability across diverse populations (Peña, 2007). The unique educational context in Sri Lanka, characterized by its distinctive linguistic and cultural characteristics, necessitates a locally adapted screening tool for effectively identifying SLDs among primary school children. These tools are crucial for addressing the particular difficulties children face in order to ensure that assessments are relevant and practical.

By reaching consensus through expert input, the Delphi method has repeatedly demonstrated its efficacy in refining the developed evaluation tools, particularly in sectors with scant empirical proof (Vogel, 2011; Keeney et al., 2011). This study not only highlights the value of the Delphi approach in refining culturally sensitive tools but also aligns with global initiatives promoting inclusive education frameworks that support children with learning disabilities, ensuring equity in educational opportunities (Kochhar and Heishman, 2010). This research contributes to the development of a relevant screening tool for early detection and intervention for children with SLDs in Sri Lanka.

The findings of this study are consistent with international efforts to develop early screening tools for learning difficulties, particularly those grounded in the identification of core cognitive and academic skill domains. For example, research by Snowling and Hulme (2020) highlights that early literacy development is strongly underpinned by phonological awareness and decoding skills, which are among the most robust predictors of later reading success; these constructs are central to widely used early assessment tools such as the Phonological Assessment Battery (PhAB) and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). Similarly, Fletcher et al. (2019) emphasise that effective screening frameworks for learning disabilities integrate both language-based and numeracy-related competencies, an approach reflected in instruments such as the Comprehensive Test of Phonological Processing (CTOPP-2) and the Test of Early Mathematics Ability (TEMA-3). These tools highlight that early deficits in number sense, symbolic processing, and phonological skills are closely associated with later academic difficulties. The prominence of these domains in the present study therefore provides empirical support for the cross-cultural relevance of these foundational constructs, while reinforcing their importance in early identification frameworks.

However, unlike many screening tools developed in Western contexts, which are often standardised within relatively homogeneous linguistic and educational systems, the present instrument incorporates linguistic, curricular, and contextual adaptations specific to Sinhala-medium learners. This distinction is critical, as prior research demonstrates that the manifestation and detection of learning difficulties are influenced by orthographic transparency, instructional practices, and sociocultural factors, which shape reading acquisition processes and the expression of learning difficulties across languages (Borleffs et al., 2019; Spencer, 2003; Katzir et al., 2012).

The study also emphasizes certain limitations. While the Delphi process facilitated the refinement of test items, field-level validation through pilot studies and broader community-based surveys remains necessary to assess the real-world efficacy of the tool and scalability, which were done

later in this study. Furthermore, the success of this screening tool depends heavily on the training of educators and their awareness of SLDs, highlighting the need for complementary capacity-building initiatives.

## Conclusion

This research which includes a Delphi survey critically highlights the importance of expert consensus in addressing the unique linguistic, cultural, and educational challenges faced by Sri Lankan primary school children. By systematically integrating expert feedback, the Delphi survey achieved a notable refinement of the initial item pool. The finalized screening tool is prepared to effectively assess SLDs, allowing early identification and interventions. The findings emphasize the critical role of culturally sensitive research methodologies in developing practical educational assessments and set a foundation for future research and interventions aimed at improving the quality of life and learning outcomes for children with SLDs in Sri Lanka.

This Delphi survey achieved its intended objective of non-statistical item reduction by systematically establishing expert consensus on the relevance, clarity, and contextual appropriateness of screening tool items for primary school children. The iterative refinement process ensured that the retained items are theoretically sound, culturally aligned, and suitable for the early identification of SLDs in Sinhala-medium learners. These findings provide a strong foundation for subsequent quantitative validation and contribute to the development of a contextually sensitive and psychometrically sound screening instrument.

## Acknowledgements

The author would like to express sincere gratitude to the panel of experts who participated in the Delphi survey and provided invaluable insights that helped refine the screening tool. The supervisors are appreciated for their invaluable guidance. Special thanks to the Ministry of Health and the Ministry of Education for the support and facilitation of this study. The contributions of the research assistants are acknowledged for their assistance in data collection and documentation. Finally, the Faculty of Graduate Studies at the University of Kelaniya is highly regarded for providing institutional support for this research.

## Declaration of Funding Sources

This part of the study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Declaration of Funding Sources

This part of the study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Data Availability Statement

The data supporting the findings of this study are not publicly available because they contain information collected from minor participants and school-based assessments that are subject to confidentiality and ethical restrictions. However, de-identified data may be obtained from the corresponding author upon reasonable request for academic and research purposes, subject to institutional approval.

## Conflict of interest statement

The authors declare no conflict of interest.

## Authorship & Contribution:

W.M.L.D.J. Wijesekara, served as the Principal Investigator, attached to the Child Care Development and Special Needs Unit, Family Health Bureau, Ministry of Health, Sri Lanka, led the overall conceptualisation, study design, implementation, data interpretation, manuscript drafting, and final coordination of the research process.

A.P.Hewamalage, contributed through academic supervision, methodological guidance, administrative facilitation, and critical review of the manuscript in the capacity of Supervisor and Unit Head of Child Care Development and Special Needs Unit, Family Health Bureau, Ministry of Health, Sri Lanka.

P.K.S. Godamunne, provided through academic supervision and senior-level academic and institutional guidance, contributed to the refinement of the research process and interpretation of findings, and critically reviewed the manuscript in the capacity of Senior Lecturer & Head of Department of Medical Education, University of Kelaniya, Sri Lanka

U.G. Karunaratne and N.H. Nanayakkara, in the capacity of research assistant attached to the Child Care Development and Special Needs Unit, Family Health Bureau, Ministry of Health, Sri Lanka, contributed substantially to the study's implementation, including data collection, coordination of research activities, documentation, support for preliminary analysis, and assistance with manuscript preparation and revisions.

## Statement

All authors accepted responsibility for the content of this paper, reviewed the findings, and approved the final version for submission.

## References

Aro, T., Eklund, K., Eloranta, A., Närhi, V., Korhonen, E., & Ahonen, T. (2018). Associations between childhood learning Disabilities and Adult-Age mental health problems, lack of

education, and unemployment. Journal of Learning Disabilities, 52(1), 71–83. https://doi.org/10.1177/0022219418775118

Barrett, C., & Heale, R. (2020). Evidence-based practice: A primer for health care professionals. Springer.

Borleffs, E., Maassen, B. A., Lyytinen, H., & Zwarts, F. (2019). Cracking the code: The impact of orthographic transparency on reading development and dyslexia. Frontiers in Psychology.

Bozatlı L, Aykutlu HC, Sivrikaya Giray A, Ataş T, Özkan Ç, Güneydaş Yıldırım B, Görker I. Children at Risk of Specific Learning Disorder: A Study on Prevalence and Risk Factors. Children (Basel). 2024 Jun 22;11(7):759.

Clark, L. A., & Watson, D. (2019). Constructing validity: Basic issues in objective scale development. Psychological Assessment, 31(12), 1412–1427.

Commonwealth of Learning. (2017). Creating learning opportunities: A toolkit for inclusive education. Commonwealth of Learning.

DeVellis, R. F. (2017). Scale development: Theory and applications (4th ed.). Sage.

DeVellis, R. F., & Thorpe, C. T. (2021). Scale development: Theory and applications (5th ed.). Sage.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2019). Learning disabilities: From identification to intervention (2nd ed.). The Guilford Press.

Garland, C. (2005). A phonological analysis of Sinhala language. UCSB Linguistics Working Papers, 17. Retrieved from https://www.linguistics.ucsb.edu/sites/secure.lsit.ucsb.edu.ling.d7/files/sitefiles/research/papers/17/garland_vol17.pdf

Grimes, P., & Cruz, J. (2021). Inclusive education in low-income countries. International Journal of Educational Development, 45(3), 1-10. Hasson, F., Keeney, S., & McKenna, H. P. (2000). Research guidelines for the Delphi survey technique. Journal of Advanced Nursing, 32(4), 1008–1015.

Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. Practical Assessment, Research & Evaluation, 12(10), 1–8.

Humphrey-Murto, S., Varpio, L., Gonsalves, C., & Wood, T. J. (2017). Using Delphi methodology in medical education research: How to do it and tips for success. Medical Teacher, 39(2), 123–129.

Katzir, T., Kim, Y. S., Wolf, M., Morris, R., Lovett, M., & Lovett, M. (2012). The impact of orthographic transparency on reading development across languages.

Keeney, S., Hasson, F., & McKenna, H. P. (2011). The Delphi technique in nursing and health research. Wiley-Blackwell.

Khodyakov, D., Grant, S., Barber, C. E., Marshall, D. A., Esdaile, J. M., Bombardier, C., & Hazlewood, G. S. (2023). Practical considerations in using online modified-Delphi approaches to engage patients and other stakeholders in clinical practice guideline development. Journal of Clinical Epidemiology, 157, 10–17.

Kochhar, C. A., & Heishman, A. C. (2010). Inclusive education: Principles and practices for effective classroom management. Pearson Education.

Lawshe, C. H. (1975). A quantitative approach to content validity. Personnel Psychology, 28(4), 563–575.

Lim, S., & Park, J.-H. (2025). Development of a questionnaire for problematic social networking sites use: Ensuring content validity through Delphi methodology. PLoS ONE, 20(3), e0315442. https://doi.org/10.1371/journal.pone.0315442

Menikdiwela, J., Rajapaksa, S., & Gunasekera, T. (2021). Challenges in identifying learning disabilities in Sri Lankan primary schools. Asian Journal of Education and Development, 5(2), 45-60.

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. Child Development, 78(4), 1255-1264.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Research in Nursing & Health, 29(5), 489–497.

Polit, D. F., & Beck, C. T. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in Nursing & Health, 30(4), 459–467.

Powell, C. (2003). The Delphi technique: Myths and realities. Journal of Advanced Nursing, 41(4), 376–382.

Redding, C. A., Maddock, J. E., & Rossi, J. S. (2006). The sequential approach to measurement of health behavior constructs: Issues in selecting and developing measures. Health Promotion, 4(1), 83–101.

Romero Jeldres, M., Díaz Costa, E., & Nadim, T. F. (2023). A review of Lawshe's method for calculating content validity in the social sciences. Frontiers in Education, 8, 1271335. https://doi.org/10.3389/feduc.2023.1271335

Scaria, L. M., Bhaskaran, D., & George, B. (2022). Prevalence of Specific Learning Disorders (SLD) among Children in India: A Systematic Review and Meta-Analysis. Indian Journal of Psychological Medicine, 45(3), 213–219. https://doi.org/10.1177/02537176221100128

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. Journal of Information Technology Education, 6, 1-21.

Snowling, M. J., Hulme, C., & Nation, K. (2011). Defining and understanding Dyslexia: Past, present and future. Oxford Review of Education, 37(4), 421-438.

Snowling, M. J., & Hulme, C. (2020). The science of reading: A handbook. Wiley.American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.; DSM-5-TR, pp. 74–77). American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787

Spencer, L. H. (2003). Effects of orthographic transparency on reading and spelling.

Vogel, R. (2011). The Delphi method in organizational research. Organizational Research Methods, 14(1), 112-138.