

INTERNATIONAL JOURNAL OF
EDUCATION, PSYCHOLOGY
AND COUNSELLING
(IJEPC)

<https://gaexcellence.com/ijepc>



CONTENT VALIDITY AND RELIABILITY OF THE FUNCTIONAL ABILITIES ASSESSMENT INSTRUMENT FOR PUPILS WITH SPECIAL EDUCATIONAL NEEDS

Nur Liyana Mohd Nazli¹, Loh Sau Cheong^{2*}, Siaw Yan-Li³

¹Faculty of Education, Universiti Malaya, Malaysia

 s2032743@siswa.um.edu.my

 <https://orcid.org/0009-0002-9309-1660>

²Faculty of Education, Universiti Malaya, Malaysia

 lohsch@um.edu.my

 <https://orcid.org/0000-0001-5962-6815>

³Faculty of Education, Universiti Malaya, Malaysia

 yanli@um.edu.my

 <https://orcid.org/0000-0002-2815-4867>

*Corresponding Author

Article Info:

Article history:

Received date: 08.01.2026

Revised date: 01.02.2026

Accepted date: 03.03.2026

Published date: 11.03.2026

To cite this document:

Nazli, N. L. M., Loh, S. C., & Siaw, Y. L. (2026). Content Validity and Reliability of The Functional Abilities Assessment Instrument for Pupils with Special Educational Needs. *International Journal of Education, Psychology and Counselling*, 11(62), 736-778.

Abstract:

Functional abilities encompass adaptive skills essential for developing independence in daily tasks, with rapid acquisition occurring during early childhood. Assessing these abilities is crucial in education programs to ensure timely interventions for learners with additional support needs. Even though international functional assessment tools are available, there are few validated instruments for Malaysian primary school pupils with special educational needs. Therefore, this study evaluates the content validity and reliability of a newly developed functional abilities assessment instrument for Malaysian primary school pupils with special educational needs. A multidisciplinary panel of experts reviewed the instrument for fairness, relevance, clarity, and ease of understanding. The Scale-Level Content Validity Index (S-CVI) results demonstrated strong agreement among experts. Some items were removed or revised based on the Item-Level Content Validity Index (I-CVI) and modified kappa coefficients (κ^*). The Rasch Measurement Model was applied across three domains to examine internal construct validity and reliability, revealing strong psychometric properties, particularly for the Body Functions and Activity and Participation sections. However, the Environmental Factors section showed slightly lower reliability indices, suggesting reduced measurement precision, though still within an acceptable range. The findings support the instrument's potential in inclusive educational settings, while highlighting the need for additional large-scale validation before broader implementation.

DOI: 10.35631/IJEPC.1162044**Keyword:**

Assessment, Content Validation, Functional Abilities, ICF-CY, Instrument, Pupils with Special Educational Needs, Rasch Analysis



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact ijepe@gaexcellence.com.

Introduction

Functional abilities refer to the practical, adaptive skills that enable individuals to manage everyday demands and participate meaningfully in daily activities. In school settings, these abilities extend beyond academic achievement to include how a child functions across cognitive, social, physical, communication, and emotional regulation domains within daily classroom and home routines (Sargent, 2025). While functional ability refers to the potential capacity of an individual to perform a task without assistance, functional performance reflects the actual execution of the task in real-world settings (Sargent, 2025).

Primary school is a critical developmental window during which children acquire and refine these foundational abilities, which directly impact their educational participation and social adjustment (Đoković et al., 2022). Pupils with special educational needs (SEN), however, may face challenges in acquiring these abilities due to difficulties in various developmental domains, which can compound over time and contribute to poor psychosocial outcomes (Horridge et al., 2016).

Despite the importance of functional abilities in supporting pupils' participation, independence, and lifelong learning, assessment practices in many educational contexts, including Malaysia, remain heavily reliant on categorical diagnostic labels when making key decisions (Frey, 2019; Xavier, 2021). Although diagnostic categorization can provide a useful starting point, it often does not yield sufficiently actionable information for teachers. Planning effective Individualized Education Plans (IEPs) requires understanding of pupils' functional profiles to guide teachers' decision-making (Klein & De Camargo, 2018).

Since the implementation of the Zero Reject Policy (ZRP) in 2019, more pupils with diverse and complex support needs have been enrolled in the national education system, making the need for robust and practical assessment instruments more pressing (Othman & Mohd Matore, 2020). Consequently, longstanding issues in service delivery, instructional differentiation, and consistent assessment of educational support needs have become increasingly urgent (Othman & Mohd Matore, 2020; Musa et al., 2021). This underscores the need for assessment instruments that are psychometrically sound, practical for regular educational use, and contextually compatible with classroom practice.

Globally, a conceptual shift has emerged to support these practical needs, in which perspectives in disability and special education have moved away from the deficit-based model toward biopsychosocial approaches that emphasize functioning and participation (Klein & De Camargo, 2018). This biopsychosocial framework was articulated in the *International Classification of Functioning, Disability and Health (ICF)* and later extended to include developmentally relevant domains in the *ICF-Children and Youth version (ICF-CY)* (WHO, 2001; WHO, 2007). Despite this shift, there remains a lack of psychometrically validated, ICF-CY-aligned functional abilities assessment instrument for Malaysian primary school pupils with special educational needs.

For this study, existing international functional and adaptive instruments were considered. However, evidence remains limited regarding their cultural transferability to non-Western cultures or low-and-middle income countries (Chiu et al., 2013; Damyanov, 2024). In the review, no instrument was identified that simultaneously met the following criteria: (1) comprehensive coverage across the full spectrum of ICF-CY domains, (2) suitability for school-based decision-making by special education teachers, and (3) cultural and linguistic fit for Malaysian primary schools. Collectively, these limitations supported the decision to develop a new locally grounded functional abilities assessment instrument rather than adopt or adapt an existing measure.

This article reports the initial phase of a broader instrument development study, focusing on the systematic construction and refinement of items that reflect functional abilities relevant to Malaysian primary school pupils with SEN and guided by the ICF-CY's domains (WHO, 2007). Specifically, the study aims to evaluate the content validity and initial psychometric properties of the developed functional abilities assessment instrument for pupils with SEN in Malaysian primary schools.

By offering an ICF-CY-referenced tool designed around local classroom practices and enabling dual reporting by parent and teacher, this study addresses a key gap in Malaysian special education assessment, which is the lack of a comprehensive instrument for profiling pupils' functional abilities in ways that support IEP goal setting and intervention planning. In addition, the study provides culturally nuanced insights from a non-Western context into the development of an ICF-CY-based assessment instrument, an area that has been predominantly shaped by Western discourse.

Literature Review

Inclusive and Special Education in Malaysia

According to Malaysia's Education Regulations (Special Education) 2013, *pupils with special needs* are those who have been diagnosed by a health professional as having a visual, hearing, speech, or physical impairment, learning disabilities, or any combination of these disabilities (Malaysia Ministry of Education, 2015). In 2022, the term was changed to *pupils with special educational needs* to make educational supports more explicit, while retaining the same definition as in the 2013 regulations (Malaysia Ministry of Education, 2022). This definition contrasts with the broader concept of SEN, which emphasizes individualized support and inclusion rather than focusing on medical classifications (Norwich, 2016).

Historically, many countries, including Malaysia, have relied on a diagnostic-to-education planning model, whereby access to special education placement and supports depended on formal medical identification (Pellicano & Houting, 2021; Bölte et al., 2024). While this approach offers procedural clarity, reliance on diagnostic labels has been criticized for limited explanatory power and poor predictive validity for educational outcomes. It also often fails to explain underlying neurological mechanisms, limiting its usefulness for planning effective supports that improve functioning and participation in schools (Werkhoven, 2022; Hollenweger, 2011).

Malaysia's special education landscape has progressively become more inclusive through key policy changes. These include the Education Act 1996, the Special Education Regulations 2013, the Malaysian Education Blueprint 2013–2025, and the Special Educational Needs Code of Practice 2014. The Zero Reject Policy, introduced in 2019, mandated that no child would be turned away from mainstream education and has become one of the most notable policy changes (Othman & Mohd Matore, 2022). Despite this progress, the identification and assessment of pupils with SEN in Malaysia remain strongly shaped by the medical model.

The need for a paradigm shifts from a deficit-based to a rights-based model of disability is becoming increasingly apparent. According to the UN Convention on the Rights of Persons with Disabilities (2006), disability should not be seen as inherent medical issues, but rather as the result of the interaction between a person's impairment and societal barriers. Within this orientation, educators are encouraged to adopt a strength-based perspective that goes beyond diagnoses by focusing on students' functional abilities.

Biopsychosocial Model of Functioning

The ICF is grounded in a biopsychosocial model that conceptualizes functioning as the outcome of dynamic interactions between health conditions, body functions and structures, activities, participation, and contextual factors (WHO, 2001). This biopsychosocial framing guides the development of the instrument, which is structured around three ICF-aligned domains: Body Functions and Structures, Activity and Participation, and Contextual Factors.

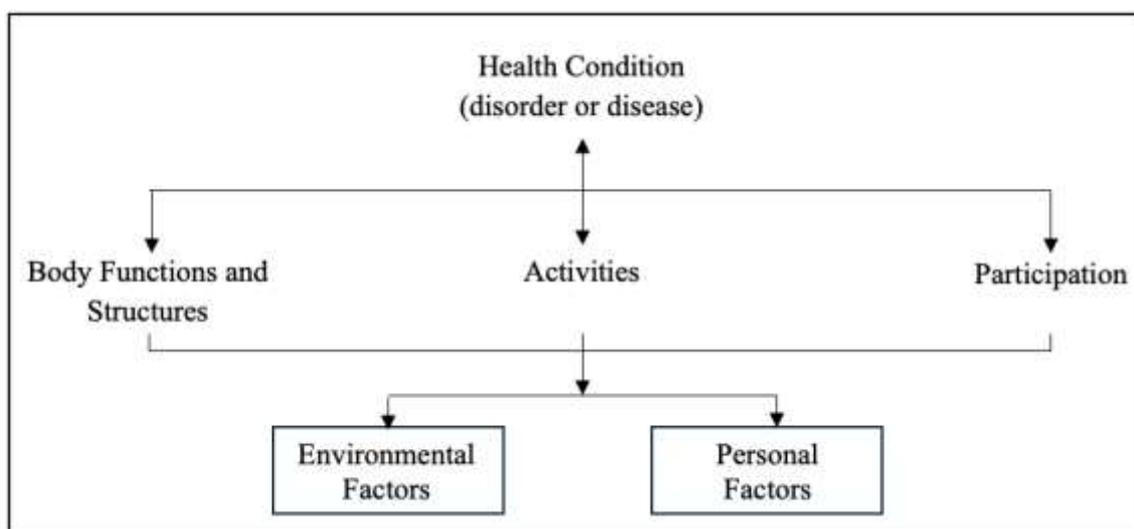


Figure 1: The Biopsychosocial Model

Source: International Classification of Functioning, Disability, and Health: ICF by the World Health Organization (WHO), 2001, World Health Organization.

Body Functions and Structures

Within the ICF, *Body Functions* refer to the physiological and psychological functions of body systems, while *Body Structures* refer to anatomical parts such as organs and limbs (WHO, 2001). A significant deviation in these functions or structures is termed an *impairment*. In educational practice, difficulties in mental functions have often been emphasized, sometimes narrowly equated with intelligence and measured via academic achievement or standardized testing. Yet, mental ability is broader and includes reasoning, planning, problem-solving, abstract thinking, and learning from experiences (Gottfredson, 1997).

Beyond cognition, a child's overall physical and mental health can influence learning-related functioning (Prokopiak & Kirenko, 2020). However, teachers may not consistently factor body-function-related impacts into educational planning because these are often perceived as clinical matters, and teachers typically have limited access to medical or diagnostic information (Silveira-Maia et al., 2019). This creates a practical gap whereby educators need structured, observable indicators of functioning that can be meaningfully interpreted for educational planning without requiring medical assessment.

Activity and Participation

In the ICF framework, *activity* refers to “the execution of a task or action by an individual,” whereas *participation* refers to “involvement in a life situation” (WHO, 2001). In simpler terms, activity is often interpreted as a person's capacity to carry out actions in a more standardized environment, while participation reflects what they actually do in their real-life environment and how they engage with the world around them (McQuaid et al., 2021). For pupils with SEN, activity and participation provide relevant lenses for education, as they capture functioning in daily learning demands and school life.

The ICF lists comprehensive domains under *Activity and Participation*, including learning and applying knowledge, general tasks and demands, communication, mobility, self-care, domestic life, interpersonal interactions and relationships, major life areas, community, social, and civic life situations (WHO, 2001). These domains are relevant to school-based functional expectations and provide a structural basis for describing pupils' functioning (Silveira-Maia et al., 2019).

Because activity and participation are culturally shaped, it is crucial that participation-based assessment tools are tailored to a community's unique cultural and local context to enable more accurate evaluation (Amini et al., 2019). When no suitable instrument exists, a conceptually equivalent measure may need to be developed so that the same construct is assessed in a way that is meaningful to the target population (Hambleton & Li, 2005).

Contextual Factors

Understanding context is crucial in profiling pupils' functioning. Urie Bronfenbrenner's Ecological Systems Theory describes how a child's development is influenced by different layers of the environment, from immediate settings like family and school (microsystem) to

broader sociocultural factors (macrosystem) (Rosa & Tudge, 2013). The ICF similarly includes *Contextual Factors* comprising *Environmental* and *Personal Factors* (WHO, 2001).

Environmental Factors include aspects of the physical, social, and attitudinal environment in which a child lives and learns (WHO, 2001). These factors can operate at the individual or societal level and can either positively or negatively influence an individual's functioning. *Personal Factors* include individual characteristics that are not part of the child's health condition, such as age, gender, temperament, and interest (WHO, 2001).

For school-aged pupils with SEN, functional performance is expressed across settings such as home and school. Therefore, meaningful educational planning requires both classroom and home perspectives: teachers observe pupils within structured instructional routines, while parents observe functioning within daily living routines in home contexts. Ecological validity can be strengthened by capturing complementary information about the pupil's functioning across these two environments (Rosa & Tudge, 2013).

International experiences indicate that applying ICF-based frameworks in education often requires localisation. The use of the ICF-CY for assessment and curriculum planning has been studied in European contexts, including Portugal, Italy, and Switzerland (Sanches-Ferreira et al., 2013; Hollenweger, 2011; Moretti et al., 2012). Similarly, Kang et al. (2025) address ICF integration in non-English education contexts, such as Taiwan and China, emphasising the need to contextualise functional ability assessment tools so they are meaningful to local educators and families.

Validity and Content Validation

In measurement, validity is defined as the accuracy of an instrument in measuring the intended construct or variable (Ahmed & Ishtiaq, 2021). At the initial phase of instrument development, the most crucial aspect is often content validity. Content validity refers to the degree to which a set of items in a scale comprehensively assesses the full theoretical scope and existing knowledge of the construct (DeVellis & Thorpe, 2022; Ahmed & Ishtiaq, 2021; Lac, 2016). Sireci and Faulkner-Bond (2014) list four elements of content validity: (1) domain definition, (2) domain representation, (3) domain relevance, and (4) appropriateness of the test construction procedure. These elements provide a framework for designing content validity evaluation.

Domain definition involves translating a complex theoretical concept into clearly defined, tangible content domains (Sireci & Faulkner-Bond, 2014). In the instrument development process, creating a theoretical blueprint can help establish content validity. In this blueprint, the content and domain areas for the measured construct are defined, and the number of items within each domain is determined (Menold et al., 2015).

Domain representation, on the other hand, refers to how accurately a test reflects and assesses the domain as outlined in its blueprint. To evaluate domain representation, a panel of subject matter experts (SMEs) is recruited to review all test items. The panel also assesses domain relevance by ensuring that the test items align appropriately with their intended domains (Sireci & Faulkner-Bond, 2014).

The literature outlines several approaches for establishing content validity, with one of the most commonly used methods being the quantification of the Content Validity Index (CVI) (Polit et al., 2007). The CVI relies on expert judgment from a panel of specialists with knowledge and expertise in the domain assessed by the instrument. These experts are asked to rate the items based on a set of evaluation criteria (Leavy, 2017).

The literature offers some guidelines on acceptable CVI values based on the number of experts on the panel. Ikart (2019) suggests engaging three to five expert reviewers for a robust content validation process, while Yusoff (2019), drawing on various sources, recommends using six to ten experts. With three to five experts, agreement needs to be very high (0.80 and above). Larger panels can accept a lower proportion of agreement, as chance agreement is less of a concern. Lynn (1986) recommended that, with six to eight experts, a CVI of 0.83 is a reasonable minimum. Similar thresholds are suggested by Polit and Beck (2006) and Polit et al. (2007).

The CVI approach was chosen instead of more complex modelling at this stage, such as the Many-Facet Rasch Model (MFRM), due to its ease of computation and interpretation, and its focus on experts' consensus regarding content relevance (Polit et al., 2007). Although the MFRM offers a more elaborate analysis, including rater bias indices and item fit statistics, it requires a much larger number of raters and items to produce stable estimates (Abu Kassim, 2024). With only a few experts, MFRM analysis would yield imprecise rater severity estimates and large standard errors for item statistics (Linacre, 1994; Boone, 2016; Azizan et al., 2020). Appropriateness of the test construction procedure is the final element of content validity. It examines the overall procedures involved in constructing a test to ensure that the instrument's content accurately and comprehensively represents the intended construct while avoiding the inclusion of irrelevant material (Sireci & Faulkner-Bond, 2014).

Reliability and Rasch Analysis

After evidence of content validity supports domain relevance and representation, reliability evidence is needed to establish that an instrument yields consistent measurement. In this study, reliability and item quality were examined using the Rasch Rating Scale Model (RRSM). Rasch analysis is widely used in scale development because it evaluates item functioning and measurement structure and produces separate reliability indices for persons and items (Bond & Fox, 2015).

Person reliability is analogous to Cronbach's alpha, but from the perspective of the reproducibility of person ordering on the trait, while item reliability indicates the reproducibility of item difficulty estimates across samples (Bond & Fox, 2015). Rasch output also provides item and person separation indices, which indicate the number of distinct levels of item difficulty and person ability that the instrument can distinguish. Desirable separation indices are typically at least 2.0, indicating that the instrument can differentiate between higher and lower levels of item difficulty and person ability (Boone, 2016).

For pilot calibration, the RRSM is feasible because it supports evaluation of item functioning even with a modest sample size, as it uses an item-response framework that can maintain stable measurement properties even with fewer respondents (Bond & Fox, 2015; Linacre, 1994). In a well-targeted polytomous test, the RRSM can still yield stable item estimates within ± 1 logit of its true value at 99% confidence with sample sizes ranging from 27 to 61 (Linacre, 1994).

Item fit statistics can inform content alignment, item clarity, and measurement precision (Boone, 2016; Bond & Fox, 2015). For polytomous data, appropriate item fit is commonly indicated by mean-square (MNSQ) outfit and infit values in the range of 0.4 to 1.5 and z-standardized (ZSTD) values within the range of -2 to +2 (Aghekyan, 2020). The RRSM also places item difficulty and person ability on the same interval scale (logits), which can be visualised using a Wright Map to evaluate item coverage and targeting (Tesio, 2003; Boone, 2016).

In addition, the RRSM can support evaluation of dimensionality and rating scale functioning. The Principal Component Analysis of Residuals (PCA-R) is used to examine whether items primarily reflect a dominant underlying construct, while rating scale diagnostics assess whether response categories operate in an ordered and interpretable manner (Fisher, 2007). Table 1 shows indicators for a quality instrument based on the Rasch analysis.

Methodology

Expert Review

The content validation process consisted of six steps, as presented in Figure 2.

Table 1: Rating Scale Instrument Quality Criteria

Criterion	Poor	Fair	Good	Very Good	Excellent
Targeting	>2 errors	1-2 errors	<1 error	<0.5 error	<0.25 error
Item Model MNSQ*	<0.33-	0.34-2.9	0.5-2.0	0.71-1.4	0.77-1.3
Range Extremes	>3.0				
Person and Item Reliability	<0.67	0.67-0.80	0.81-0.90	0.91-0.94	>0.94
Cronbach's Alpha (KR20)	<0.5	0.5-0.6	0.6-0.7	0.7-0.8	>0.8
Unexplained variance in contrast 1-5 of PCAR**	>15%	10-15%	5-10%	3-5%	<3%

Note: *MNSQ=Mean Square; **PCAR=Principal Component Analysis of Residuals.

Source: Rating Scale Instrument Quality Criteria by William P Fisher Jr, 2007.

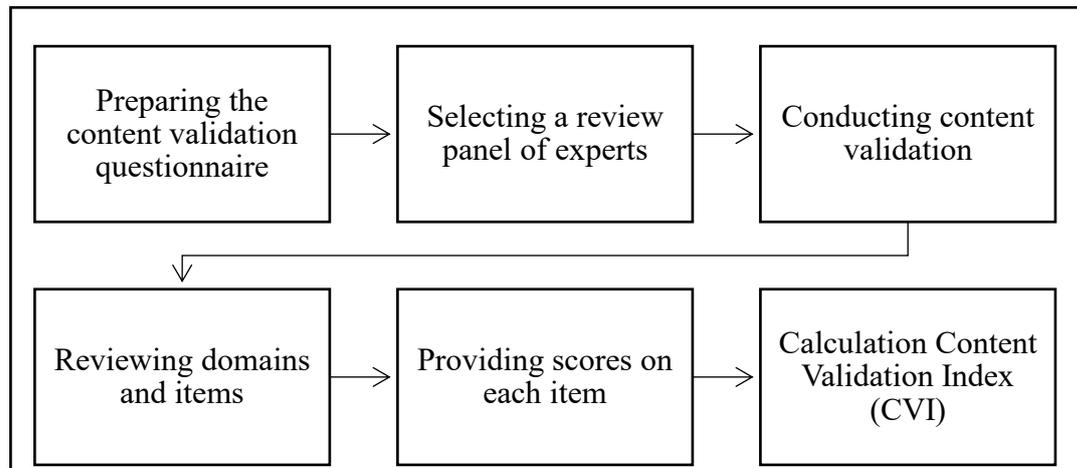


Figure 2: Determining Content Validity Process

Source: ABC of Content Validation and Content Validity Index Calculation Muhamad Saiful Bahri Yusoff, 2019, Education in Medicine Journal.

This study computed two types of Content Validity Index (CVI): item-level CVI (I-CVI) and scale-level CVI (S-CVI). Prior to these calculations, the rating scale was simplified into two dichotomous categories, with a rating of 1 classified as "content invalid" and ratings of 2 and 3 considered "content valid." This study calculated S-CVI using the average method, as the universal method was too stringent when involving many experts (Polit & Beck, 2006).

The study used multi-rater kappa statistics to complement the CVI and to account for chance agreement. The chance agreement can be an issue when choices are dichotomous, as in this content validation (Polit et al., 2007). Following the recommendations of Landis and Koch (1977) and Shrout (1998), the acceptable range of modified Cohen's kappa (κ^*) values is between 0.81 and 1.00.

Based on published guidance, eight experts were engaged because this panel size falls within the recommended range for CVI-based content validation (six to ten experts), as suggested by Yusoff (2019). This number also aligns with Lynn's (1986) recommendation that a minimum CVI value of 0.83 is appropriate when six to eight experts are used. This panel size reduces the influence of chance agreement compared with smaller panels and allows adequate multidisciplinary representation (Polit & Beck, 2006; Polit et al., 2007; Lynn, 1986).

The inclusion criteria for selecting the expert panel were carefully established to ensure comprehensive, multidisciplinary input in the instrument validation process. Experts were selected based on their qualifications, professional experience, and relevance to the domains being assessed. All experts selected had between 5 and 10 years of experience in their respective fields, ensuring a well-rounded and credible validation process. Details of the recruited experts are presented in Table 2.

Table 2: The Panel of Experts for Expert Review

Expert	Field of Expertise	Highest Qualification	Professional Practice	Years of Experience
Expert 1	Occupational Therapy	PhD in Occupational Therapy	Assistant Director (Occupational Therapy), Ministry of Education.	13
Expert 2	Clinical Psychology	PhD of Health and Behavioural Science	Assistant Professor, and Clinical Psychologist	10
Expert 3	Care and Community	PhD in Management	Chief Assistant Director, Department of Development for Persons with Disabilities, Department of Social Welfare (JKM)	25
Expert 4	Education (Ministry of Education)	PhD (Special Education Management)	Professor, and Ex- Director, Special Education Division, Ministry of Education	40
Expert 5	Instrument Development	PhD (Ergonomics)	Associate Professor and Dean	24
Expert 6	Language	PhD (Psycholinguistic)	Senior Lecturer	22
Expert 7	Education (Special Education teacher)	Bachelor of Education (Special Education)	School Inspector for State Education Department, and ex-special education teacher.	29
Expert 8	Parent of Pupils with Special Needs	Bachelor of Engineering (Electrical)	Fulltime Housewife	15

Instrument

The researcher provided the expert panel with a content validation form designed to systematically gather their evaluations. The form included a briefing page, the criteria and rating scale, expert's demographic details of the experts, and a list of items accompanied by the rating scale.

An initial pool of 173 items was generated from the ICF-CY 7-12-year code set and was written in Malay (Ellingsen & Simeonsson, 2024). The item pool was deliberately inclusive to ensure comprehensive sampling across ICF-CY domains and to reduce the risk of construct underrepresentation at the early stage of development. The items were assessed using four specific criteria: fairness, relevance, clarity, and ease of understanding.

Although 4-point Likert scales are commonly used in content validity assessments, they have limitations, particularly the potential to inflate inter-rater agreement due to chance. To address this issue, Almanasreh (2020) recommended using either a 5-point or 3-point scale, as these

options may offer greater reliability. This study opted for a 3-point rating scale to minimize cognitive overload for experts tasked with evaluating multiple items.

Pilot Study

The pilot study sample consisted of 30 pupils with SEN from the Special Education Integration Program (SEIP) in Malaysian primary schools. Drawing on a synthesis of the literature on pilot-study sample sizes, Johanson and Brooks (2010) recommend that around 30 participants represent a reasonable minimum for pilot testing in initial scale development. Given that stable item estimates (± 1 logit at 99% confidence) can be achieved in well-targeted polytomous measures with samples as small as 27 respondents, a sample of 30 pupils was deemed adequate for this pilot calibration. This decision to limit the reliability test sample to 30 pupils was also guided by practical and ethical considerations, as the study involved a sensitive population.

The pilot study was conducted at two schools in the Hulu Langat district of Selangor. Selangor was chosen due to its high population of pupils with special educational needs, with Hulu Langat having the highest concentration within the state. This selection helped to ensure a well-represented sample that captured a diverse population. Table 3 presents the demographic data on the pupil participants ($n=30$) involved in the pilot study.

Table 3: Demographic Data of Pupil Participants in the Pilot Study

Demographic Data	Frequency (n)	Percentage (%)
Gender		
Male	19	63.3
Female	11	36.7
Age (M = 8;06)		
6;0-7;11	13	43.3
8;0-9;11	12	40.0
10;0-11;11	4	13.3
12;0-13;11	1	3.3
Disability Category		
Dyslexia	4	13.3
Autism Spectrum Disorder	17	56.7
Vision Impairment	2	6.7
Multiple Disabilities	3	10.0
ADHD	1	3.3
Intellectual Disabilities	1	3.3
Others	2	6.7

Once the study had received approval from the university's ethics committee, permission was obtained from the Ministry of Education and the State Education Department to conduct the study in schools. After consent was received from school administrators, the researcher briefed teachers about the study and provided training on using the instrument. Due to educational research policy, the instrument had to be administered by teachers as proxies.

Twelve teachers (n=12) were involved in the pilot study. All were female and held at least a bachelor's degree. Their mean age was 39.83 years, and the mean years of experience was 12.67, indicating a well-seasoned cohort. Each teacher was required to administer the instrument to two to three pupils under their care. They were given two to three months to conduct observations and complete the instrument. After the instruments were completed, the researcher collected the documents for analysis. The instrument data were analysed using WINSTEPS version 5.8.1 and examined separately in three sections: (1) Body Functions, (2) Activity and Participation, and (3) Environmental Factors.

Instrument

The instrument consisted of five main sections as presented in Table 4. These sections comprise subdomains made up of clusters of items. The instrument also includes a page for recording pupils' demographic information, and another page summarizing scores for each section.

Table 4: Sections and Number of Items

No	Domain	No. of Items	
		Original Items	After Expert Review
1	Body Function	44	49
2	Activity and Participation at School	64	95
3	Activity and Participation at Home	45	47
4	Environmental Factors at School	9	12
5	Environmental Factors at School	11	14
TOTAL NUMBER OF ITEMS		173	217

The instrument uses a 4-point rating scale across sections, with 0=Not applicable/Unknown, which is used when the teacher or parent has insufficient information about an item. For Body Functions, the scale rates how often the pupil demonstrates the ability (1=Not at all, 2=Sometimes, 3=Always). For Activity and Participation, the scale rates the level of assistance the pupil needs to perform the activity (1=Full assistance, 2=Partial assistance, 3=No assistance needed). For Environmental Factors, the scale indicates how the factor affects participation at school or home (1=Hinders participation, 2=Neutral, 3=Facilitates participation).

Results

Content Validity

Table 5 outlines the S-CVI values for the four criteria: fairness, relevance, clarity, and ease of understanding. All criteria show values greater than 0.90, surpassing the acceptable threshold of 0.83.

Table 5: S-CVI Values

Criteria	Fairness	Relevance	Clarity	Ease of Understanding
S-CVI	0.97	0.97	0.98	0.94

Table 6: Summary of Items Requiring Attention

Domain	Item	I-CVI/Modified Kappa Value				Comment	
		Fairness	Relevance	Clarity	Easy		
BODY FUNCTION							
1	Cognitive Function	1.1	0.88/ 0.87	0.75/ 0.72	0.75/ 0.72	1.00/ 1.00	The item is too general.
		1.3	1.00/ 1.00	1.00/ 1.00	1.00/ 1.00	0.75/ 0.72	This item is considered difficult to answer.
5	Other Body Function	5.1	0.88/ 0.87	0.88/ 0.87	0.88/ 0.87	0.75/ 0.72	These items were questioned for its relevance for teachers to fill in as teachers may not have knowledge about these aspects.
		5.2	0.88/ 0.87	0.88/ 0.87	0.88/ 0.87	0.75/ 0.72	
		5.3	0.75/ 0.72	0.88/ 0.87	0.88/ 0.87	0.75/ 0.72	
		5.4	0.75/ 0.72	0.88/ 0.87	0.88/ 0.87	0.75/ 0.72	
ACTIVITY AND PARTICIPATION							
2	Basic Writing Skills	2.2	0.75/ 0.72	0.88/ 0.87	0.88/ 0.87	0.75/ 0.72	Double-barrelled item.
		2.3	0.75/ 0.72	0.75/ 0.72	0.88/ 0.87	0.75/ 0.72	Double-barrelled item.
		2.6	0.88/ 0.87	0.75/ 0.72	1.00/ 1.00	0.88/ 0.87	Consider not relevant for SEN pupils
3	Basic Counting Skills	3.5	0.88/ 0.87	0.63/ 0.68	1.00/ 1.00	0.75/ 0.72	Consider not relevant for SEN pupils
5	Thinking Skills	5.3	0.75/ 0.72	0.75/ 0.72	0.75/ 0.72	0.63/ 0.68	Consider not relevant for SEN pupils
6	Problem Solving	6.2	0.75/ 0.72	0.75/ 0.72	0.75/ 0.72	0.75/ 0.72	Consider not relevant for SEN pupils
7	Completing Task	7.4	0.75/ 0.72	0.75/ 0.72	0.88/ 0.87	0.75/ 0.72	Consider not relevant for SEN pupils
13	Body Care	13.1	0.75/ 0.72	0.75/ 0.72	0.75/ 0.72	0.75/ 0.72	Need to review the item clarity and relevance
ENVIRONMENTAL FACTORS							
2	At Home	1.8.1	0.88/ 0.87	0.88/ 0.87	0.88/ 0.87	0.75/ 0.72	The item may not be applicable to all pupils.

The I-CVI values for all items ranged from 0.63 to 1.00. The modified kappa coefficients (κ^*) for all items ranged from 0.68 to 1.00. The detailed I-CVI and κ^* values for all items are listed in Appendix. Some items fell below the acceptable I-CVI threshold and the κ^* range. Table 6

summarizes the items that fell below the CVI value of 0.83 and outside the acceptable κ^* range of 0.81 to 1.00.

A total of 15 items were identified with low I-CVI or κ^* values across one or more evaluation criteria, comprising six items from the Body Functions section, eight items from the Activity and Participation section, and one item from the Environmental Factors section. These items were distributed across nine subdomains.

These findings were cross-checked against the expert panel's qualitative comments to clarify the issues associated with each flagged item. Seven main issues identified: (1) overly general items, (2) inappropriate levels of difficulty, (3) limited relevance to teachers' assessment, (4) limited relevance to pupils with special educational needs, (5) lack of clarity, (6) double-barrelled items, and (7) limited applicability to all pupils within the population. The flagged items were re-evaluated for suitability and fairness in measuring the intended constructs. They were also reviewed for sentence structure and wording to ensure consistent interpretation and responses across respondents.

Reliability

The three sections of the instrument were analysed separately, as they measure distinct constructs using section-specific rating scales. Overall, the Body Functions and Activity and Participation sections demonstrate adequate Rasch measurement precision, with reliability and separation indices meeting commonly used criteria. The results indicate that the items can be replicated across samples and that the scales can differentiate pupils across multiple ability strata.

In contrast, the Environmental Factors section shows reduced measurement precision, with person and item reliability and separation indices below recommended thresholds. This finding suggests a limited ability to distinguish pupils and item difficulty levels in this pilot sample. Although internal consistency is acceptable and overall variance explained meets the minimum guideline, the separation results indicate that further refinement and re-testing are needed to improve discrimination and stability. Table 7 summarises the results of this reliability testing.

Table 7: Reliability of the Instrument's Sections

Statistics	Threshold	Reliability Results for Each Section		
		Body Functions	Activity and Participation	Environmental Factors
Person reliability	>0.80	0.96	0.99	0.69
Item reliability (real)	>0.80	0.89	0.93	0.61
Cronbach's Alpha	>0.80	0.98	0.99	0.93
Person separation index	>2.00	4.88	9.43	1.14
Item separation index	>2.00	2.80	3.61	1.25
Raw variance explained by measures	>40%	59.9%	66.2%	42.8%
Unexplained variance in the first contrast	<15%	7.4%	5.3%	8.9%
Eigenvalue of the first contrast	<3.00	9.01	22.18	4.03

PCA of standardized residuals supports the assumption of unidimensionality for this instrument. The raw variance explained by the measures across the three sections exceeded the minimum threshold of 40%, which indicates that a substantial proportion of the variance in item responses can be attributed to the latent trait being measured. Despite the eigenvalues of the first contrast exceeding the acceptable threshold of 3.0 in all domains, the unexplained variance in the first contrast remained well below the maximum recommended value of 15%. This indicates that no significant secondary dimension was present in the residuals.

Discussion

Content Validity

Ensuring content validity is a crucial step before evaluating an instrument's construct validity (Haynes et al., 1995). The primary aim of content validity is to verify that the instrument contains a sufficient and relevant set of items that accurately measure the intended construct (Polit et al., 2007).

Content validity is evidenced by the excellent scale-level content validity index (S-CVI) values for fairness, relevance, clarity, and ease of understanding, all of which exceed the acceptable threshold of 0.83. These results reflect strong agreement among the expert panel regarding the instrument's quality (Polit & Beck, 2006). The fairness and relevance S-CVI values were 0.97, indicating that the items in the instrument are relevant and unbiased in measuring pupils' functional abilities, regardless of their race, religion, type, or level of disability.

The high S-CVI values for clarity and ease of understanding also indicate that the items are written clearly, without ambiguity in meaning. These qualities are essential to ensure that respondents, regardless of their academic background or proficiency level, can comprehend and interpret the items consistently. This aligns with Brosnan et al. (2021), who recommend on strategies to reduce task difficulty when participants respond to an instrument, including keeping it concise, avoiding jargon, and using clear, simple sentences.

At the item level (I-CVI), fifteen items were found to fall below the acceptable threshold of 0.83, with lower modified kappa (κ^*) values. These results raise concerns regarding the fairness, relevance, clarity, and ease of understanding of these items. The items were reviewed and modified accordingly. Following the expert review, the number of items increased from the initial item pool of 173 to 217. Although some irrelevant items were omitted based on expert feedback, the increase occurred because several items were elaborated or broken down to avoid double-barrelled constructions.

In the Body Functions section, the 'Other Body Functions' domain, which assesses general health aspects such as skin condition, immune system performance, cardiovascular functions, and respiratory functions, received overall criticism from the expert panel. These items were considered outside teachers' observational capacity, as they require clinical evaluation and access to medical records. Such items can introduce bias, as raters who are unfamiliar with observing and assessing clinically oriented content may provide inaccurate interpretations (Luoto et al., 2023). Omitting these items helps ensure that the instrument remains accessible and appropriate for non-clinical raters, such as school teachers (Ilium & Gradel, 2017).

In the Activity and Participation section, several items measuring functional academics, such as reading, writing, and counting skills, were identified as double-barrelled questions (DBQs) or overly broad in scope. DBQs require respondents to address multiple aspects within a single response, which increases cognitive load and extends response times, thereby compromising the efficiency and accuracy of the assessment (Menold, 2020).

Some items were further delineated to align with established literacy, numeracy, and writing skill frameworks. This alignment facilitated a structured progression of tasks, moving from simple to complex, thereby more accurately capturing stages of skill development. Consequently, additional items were incorporated into the instrument. This is a common practice in psychometric assessment, whereby a construct is initially defined in broad terms and subsequently refined into more concrete, measurable components to enhance both reliability and validity in educational settings (DeVellis & Thorpe, 2022).

There were not many issues with the Environmental Factors section, except for one item from the Home Environment domain, which was flagged as inapplicable to all pupils. Based on these findings, modifications were implemented to improve the clarity and relevance of the affected items. Some items were refined for wording, while others were removed following expert feedback. These adjustments enhance the instrument's overall validity by ensuring alignment with the intended assessment framework.

Reliability

The reliability analysis of the functional abilities assessment instrument demonstrated strong measurement performance, particularly in the Body Functions and Activity and Participation sections. Both sections exhibited high Cronbach's alpha values (>0.98), along with person and item reliability indices exceeding 0.89 and strong separation indices (Bond & Fox, 2015). These results indicate that the instrument provides consistent measurement of functional abilities within these domains.

However, the Environmental Factors section exhibited lower reliability. Although Cronbach's alpha remained above the acceptable threshold (>0.8), the person and item reliability indices fell below 0.80 (Bond & Fox, 2015; Sumintono & Widhiarso, 2015). These results indicate reduced measurement precision. In addition, the person and item separation indices suggest that the domain does not effectively differentiate items and persons into more than two levels of difficulty and ability.

This pilot study involved a small sample size of 30 pupils with special educational needs. While this is adequate for exploratory analyses, the small sample limits the stability of parameter estimates to approximately ± 1 logit with 99% confidence (Linacre, 1994). A small sample size of this magnitude increases the root mean square error (RMSE) and may introduce bias in item parameter estimates (Azizan et al., 2020). Although this was not a concern for the Body Functions (49 items) and Activity and Participation (142 items) sections, the smaller number of items in the Environmental Factors section (26 items) may have affected the stability of the estimates.

However, Sumintono and Widhiarso (2015) note that reliability indices greater than 0.67 can be considered sufficient in the early stages of instrument development. Hence, the Environmental Factors section was retained for further evaluation in the main study with a

larger sample. The results also highlight that relying on Cronbach's alpha alone is not recommended and that it should be interpreted alongside other reliability indices for a more robust evaluation (Tavakol & Dennick, 2011).

To facilitate a direct comparison between person and item characteristics while keeping them distinct, the Rasch model requires two fundamental conditions: (1) unidimensionality and (2) local independence (Abu Kassim, 2024). Based on the reliability findings, the dimensionality analysis showed that the raw variance explained by the measures exceeded 40% for all sections, supporting the instrument's structural validity. The unexplained variance in the first contrast across all sections was also below 15%, indicating that the influence of a second dimension is minimal.

These findings support the Rasch model's fundamental requirement of unidimensionality, confirming that the assessment effectively provides a coherent and focused measurement of functional abilities (Bond & Fox, 2015). Furthermore, the minimal influence of a second dimension suggests that the instrument is not substantially affected by construct contamination, reinforcing its validity.

Implications of the Study

The development of the instrument to assess functional abilities of pupils with special educational needs in Malaysian primary schools has important implications for policy, practice, methodology, and theory in Malaysian special and inclusive education. By providing a structured, strength-based, ICF-CY grounded instrument, it directly addresses a key implementation gap in the Zero Reject Policy and Classroom-Based Assessment (PBD). The instrument promotes equitable support planning, which is often constrained by the lack of context sensitive assessment tools. Its accessible language and multi-informant design support meaningful parent-teacher collaboration in IEP goal setting and intervention planning by capturing home and school contexts.

Methodologically, the study demonstrates a systematic multi-stage validation approach, supporting evidence-informed assessment that is fit for purpose, fair, and interpretable for educational decision making. Theoretically, the instrument advances the shift from diagnosis-led classification to a biopsychosocial, rights-based approach of disability. This study contributes to broader international discourse on ICF-CY-informed functional assessment in non-Western settings by providing locally relevant empirical evidence.

Limitations and Future Studies

This study is part of a broader instrument development process aimed at creating a standardized assessment of functional abilities for pupils with SEN in Malaysian primary schools. Despite the promising outcomes, the study has several limitations that should be addressed in the subsequent phase or in future studies.

The small sample size of pupils with SEN limits the generalizability of the findings to a larger population. A larger validation study is needed to obtain more conclusive and generalizable evidence regarding the instrument's effectiveness in measuring functional abilities across a more diverse population (Azizan et al., 2020). While trait homogeneity is necessary in Rasch analysis, ability homogeneity can reduce the effectiveness of Rasch calibration (Linacre, 1994).

Including respondents from more diverse demographic backgrounds may improve estimate stability for the Environmental Factors section.

Although the instrument is highly structured, it may be influenced by raters' subjective judgements. Uncontrolled variables such as teachers' training, teaching experience, and differing interpretations of items may introduce bias. Using the Many-Facet Rasch Model (MFRM) can help evaluate data when multiple raters are involved by incorporating additional facets beyond persons and items (Abu Kassim, 2024).

A comprehensive functional abilities assessment considers a child's adaptive functioning across three settings: school, home, and the community (WHO, 2001). This current instrument measures functioning only in school and home settings and does not include many community-based activities. Given the important role of community participation in a child's development, the absence of a community component is a gap that needs to be addressed in future studies.

With a larger sample in the subsequent phase, the instrument can benefit from Rasch-based refinement to produce a more efficient and psychometrically sound version. The current version is lengthy, comprising 217 items. Redundant items can be identified through overlapping logits and evidence of local dependence and then removed to create a shorter version that improves precision and reduces response burden for teachers (Bond & Fox, 2015).

Conclusion

Expert content validation using the CVI and Rasch Measurement Model helped refine the instrument and align it ICF-CY-based constructs. The S-CVI supported overall conceptual alignment, while I-CVI ratings confirmed item relevance, clarity, and representativeness; items below the threshold were revised or removed based on expert feedback. Rasch analysis then provided empirical evidence of item functioning, reliability, separation, and domain one-dimensionality, with the Body Functions and Activity and Participation sections meeting recommended thresholds. Although the Environmental Factors section showed slightly lower reliability, it remained acceptable for early-stage development and warrants further testing. Overall, this dual approach strengthened the instrument's validity and theoretical grounding. Future studies should validate the instrument with larger, more diverse samples across settings and disability categories to support broader standardization and potential use alongside diagnostic assessments for improved multidisciplinary planning.

-
- Acknowledgements:** The authors gratefully acknowledge Universiti Malaya and the Ministry of Education Malaysia for the resources and support provided throughout this research. Special appreciation is extended to colleagues and peers for their valuable insights and constructive feedback, which strengthened the quality of the paper.
- Funding Statement:** No Funding
- Conflict of Interest Statement:** The authors declare that there is no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the International Journal of Education, Psychology and Counseling (IJEPC).
- Ethics Statement:** This study was conducted in accordance with ethical research standards. All procedures involving human participants were reviewed and approved by the Universiti Malaya Research Ethics Committee (UMREC), approval number UM.TNC2/UMREC_3182. Informed consent was obtained from all participants prior to data collection. Participation was voluntary, and respondents were assured of confidentiality and anonymity. The data collected were used solely for academic purposes.
- Author Contribution Statement:** All authors contributed substantially and equally to the development of this manuscript. Nur Liyana Mohd Nazli led the data collection, conducted the analysis, and interpreted the results, and contributed to drafting the manuscript. Loh Sau Cheong provided overall supervision of the study, contributed to the conceptualization and methodology, and critically revised the manuscript. Siaw Yan-Li co-supervised the study, contributed to the conceptualization and methodology, and critically revised the manuscript. All authors reviewed and approved the final version of the manuscript prior to submission.
-

References

- Abu Kassim, N. L. (2024). The RASCH model for psychometric testing rater-mediated assessment and standard setting. IIUM Press.
- Aghekyan, R. (2020). Validation of the SIEVEA instrument using the Rasch analysis. *International Journal of Educational Research*, 103, 101619. <https://doi.org/10.1016/j.ijer.2020.101619>
- Ahmed, I., & Ishtiaq, S. (2021). Reliability and validity: Importance in Medical Research. *PubMed*, 71(10), 2401–2406. <https://doi.org/10.47391/jpma.06-861>
- Almanasreh, E., Moles, R., & Chen, T. F. (2018). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, 15(2), 214–221. <https://doi.org/10.1016/j.sapharm.2018.03.066>
- Amini, M., Mehraban, A. H., Pashmdarfard, M., & Cheraghifard, M. (2019). Reliability and validity of the Children Participation Assessment Scale in Activities Outside of School–Parent version for children with physical disabilities. *Australian Occupational Therapy Journal*, 66(4), 482–489. <https://doi.org/10.1111/1440-1630.12569>
- Azizan, N. H., Mahmud, Z., & Rambli, A. (2020). Rasch Rating Scale Item Estimates using Maximum Likelihood Approach: Effects of Sample Size on the Accuracy and Bias of the Estimates. *International Journal of Advanced Science and Technology*, 29(4s), 2526–2531.
- Bölte, S., Alehagen, L., Black, M. H., Hasslinger, J., Wessman, E., Remnélius, K. L., Marschik, P. B., D’Arcy, E., Crowson, S., Freeth, M., Seidel, A., Girdler, S., & Zander, E. (2024). The Gestalt of functioning in autism revisited: First revision of the International Classification of Functioning, Disability and Health Core Sets. *Autism*, 28(9), 2394–2411. <https://doi.org/10.1177/13623613241228896>
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge/Taylor & Francis Group.
- Boone, W. J. (2016). Rasch Analysis for Instrument Development: Why, when, and how? *CBE—Life Sciences Education*, 15(4), rm4. <https://doi.org/10.1187/cbe.16-04-0148>
- Brosnan, K., Grün, B., & Dolnicar, S. (2021). Cognitive load reduction strategies in questionnaire design. *International Journal of Market Research*, 63(2), 125–133. <https://doi.org/10.1177/1470785320986797>
- Chiu, W., Yen, C., Teng, S., Liao, H., Chang, K., Chi, W., Wang, Y., & Liou, T. (2013). Implementing disability evaluation and welfare services based on the framework of the international classification of functioning, disability and health: experiences in Taiwan. *BMC Health Services Research*, 13(1). <https://doi.org/10.1186/1472-6963-13-416>
- Damyantov, K. (2024). International Classification of Functioning, Disability, and Health for Children and Youth (ICF-CY) as a Framework for Early Intervention Planning in Inclusive Education. *Open Journal of Social Sciences*, 12(11), 85–98. <https://doi.org/10.4236/jss.2024.1211006>
- Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5(4), 194–197. [https://doi.org/10.1016/s0897-1897\(05\)80008-4](https://doi.org/10.1016/s0897-1897(05)80008-4)
- DeVellis, R. F., & Thorpe, C. T. (2022). *Scale Development: theory and applications* (5th ed.). SAGE.
- Doković, S., Kovačević, T., & Ostojić-Zeljkić, S. (2022). Functional abilities of children as factors of school achievement. *Zbornik Radova Filozofskog Fakulteta U Pristini*, 52(3), 277–297. <https://doi.org/10.5937/zrffp52-35020>

- Ellingsen, K. M., & Simeonsson, R. J. (2024). Defining childhood disability: ICF-CY developmental code sets. *Disability and Rehabilitation*, 1–9. <https://doi.org/10.1080/09638288.2024.2303385>
- Fisher Jr., W. P. (2007). Rating Scale Instrument Quality Criteria. *Rasch Measurement Transaction*, 21, 1095. <http://www.rasch.org/rmt/rmt211a.htm>
- Frey, J. R. (2019). Assessment for special education: diagnosis and placement. *The Annals of the American Academy of Political and Social Science*, 683(1), 149–161. <https://doi.org/10.1177/0002716219841352>
- Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1), 13–23. [https://doi.org/10.1016/s0160-2896\(97\)90011-8](https://doi.org/10.1016/s0160-2896(97)90011-8)
- Hambleton, R. K., & Li, S. (2005). Translation and adaptation issues and methods for educational and psychological tests. In *Comprehensive Handbook of Multicultural School Psychology* (pp. 881–903). John Wiley & Sons, Inc.
- Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3), 238–247. <https://doi.org/10.1037/1040-3590.7.3.238>
- Hollenweger, J. (2011). Development of an ICF-based eligibility procedure for education in Switzerland. *BMC Public Health*, 11(S4). <https://doi.org/10.1186/1471-2458-11-s4-s7>
- Horridge, K. A., McGarry, K., Williams, J., & Whitlingum, G. (2016). Prospective pilots of routine data capture by paediatricians in clinics and validation of the Disabilities Complexity Scale. *Developmental Medicine & Child Neurology*, 58(6), 581–588. <https://doi.org/10.1111/dmcn.13101>
- Ikart, E. M. (2019). Survey questionnaire Survey pretesting method: an evaluation of survey questionnaire via expert reviews technique. *Asian Journal of Social Science Studies*, 4(2), 1. <https://doi.org/10.20849/ajsss.v4i2.565>
- Johanson, G. and Brooks, G. (2010) Initial Scale Development: Sample Size for Pilot Studies. *Educational and Psychological Measurement*, 70, 394-400. <http://dx.doi.org/10.1177/0013164409355692>
- Kang, L., Naude, A., Rocha, A., Bornman, J., Alves, I., Maxwell, G. R., Moretti, M., Proyer, M., Shi, L., & Zahnd, R. (2025). An international systematic review of the ICF in education: cross-language comparisons from Chinese, German, Italian, and Portuguese publications. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.1526194>
- Klein, B., & De Camargo, O. K. (2018). A proposed functional abilities classification tool for developmental disorders affecting learning and behaviour. *Frontiers in Education*, 3. <https://doi.org/10.3389/feduc.2018.00002>
- Lac, A. (2016). Content analysis. In R. J. R. Levesque (Ed.), *Encyclopedia of Adolescence* (2nd ed., pp. 1-5). Switzerland: Springer International.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Leavy, P. (2017). *Research design: Quantitative, qualitative, mixed methods, Arts-Based, and Community-Based participatory research approaches*. The Guilford Press.
- Linacre, J. M. (1994). Sample size and item calibration or person measure stability. *Rasch*. <https://www.rasch.org/rmt/rmt74m.htm>
- Luoto, J., Klette, K., & Blikstad-Balas, M. (2022). Possible biases in observation systems when applied across contexts: conceptualizing, operationalizing, and sequencing instructional quality. *Educational Assessment Evaluation and Accountability*, 35(1), 105–128. <https://doi.org/10.1007/s11092-022-09394-y>

- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382-386. <https://doi.org/10.1097/00006199-198611000-00017>
- Malaysian Ministry of Education. (2015). *Buku Panduan Pengoperasian PPKI. Bahagian Pendidikan Khas.*
- Malaysian Ministry of Education. (2022). *Malaysian Education Blueprint 2013-2025: Annual Report 2022.* In Ministry of Education. <https://www.moe.gov.my/storage/files/shares/Dasar/PPPM/MEB%20Annual%20Report%202022.pdf>
- McQuaid, G. A., Pelphrey, K. A., Bookheimer, S. Y., Dapretto, M., Webb, S. J., Bernier, R. A., McPartland, J. C., Van Horn, J. D., & Wallace, G. L. (2021). The gap between IQ and adaptive functioning in autism spectrum disorder: Disentangling diagnostic and sex differences. *Autism*, 25(6), <https://doi.org/10.1177/1362361321995620>
- Menold, J., Jablokow, K., Purzer, S., Ferguson, D., & Ohland, M. (2015). Using an instrument blueprint to support the rigorous development of new surveys and assessments in engineering education. 122nd ASEE Annual Conference & Exposition, 11853. <https://doi.org/10.18260/p.24993>
- Menold, N. (2020). Double Barreled Questions: An analysis of the similarity of elements and effects on measurement quality. *Journal of Official Statistics*, 36(4), 855–886. <https://doi.org/10.2478/jos-2020-0041>
- Moretti, M., Alves, I., & Maxwell, G. (2012). A Systematic Literature review of the situation of the International Classification of Functioning, Disability, and Health And the International Classification of Functioning, Disability, and Health–Children and Youth Version in Education. *American Journal of Physical Medicine & Rehabilitation*, 91(13), S103–S117. <https://doi.org/10.1097/phm.0b013e31823d53b2>
- Musa, S. H., Md Zain, E., Ismail, M. Z., Mat Hussin, H., & Nordin, M. N. (2021). Something important for special education in Malaysia. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(11), 5211–5216. <https://doi.org/10.17762/turcomat.v12i11.6733>
- Norwich, B. (2016). Conceptualizing Special Educational Needs using a Biopsychosocial Model in England: The Prospects and Challenges of using the International Classification of Functioning Framework. *Frontiers in Education*, 1. <https://doi.org/10.3389/feduc.2016.00005>
- Othman, N., & Mohd Matore, M. E. E. (2020). The Zero-Reject Policy in Special Education: A critique review. *Journal of Critical Reviews*, 7(11), 2507–2515.
- Pellicano, E., & Houting, J. D. (2021). Annual Research Review: Shifting from ‘normal science’ to neurodiversity in autism science. *Journal of Child Psychology and Psychiatry*, 63(4), 381–396. <https://doi.org/10.1111/jcpp.13534>
- 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>
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4), 459–467. <https://doi.org/10.1002/nur.20199>
- Prokopiak, A., & Kirenko, J. (2020). ABAS-3 – an instrument for assessing adaptive skills in people with an intellectual disability. *Hrvatska Revija Za Rehabilitacijska Istraživanja*, 56(2), 154–168. <https://doi.org/10.31299/hrri.56.2.9>
- Rosa, E. M., & Tudge, J. (2013). Urie Bronfenbrenner’s Theory of Human Development: Its Evolution From Ecology to Bioecology. *Journal of Family Theory & Review*, 5(4), 243–258. <https://doi.org/10.1111/jftr.12022>

- Sanches-Ferreira, M., Lopes-Dos-Santos, P., Alves, S., Santos, M., & Silveira-Maia, M. (2013). How individualised are the Individualised Education Programmes (IEPs): an analysis of the contents and quality of the IEPs goals. *European Journal of Special Needs Education*, 28(4), 507–520. <https://doi.org/10.1080/08856257.2013.830435>
- Sargent, L. (2025). Functional ability [E-Book]. In *Concepts for Nursing Practice* (4th ed., pp. 14–22). Elsevier.
- Shrout, P. E. (1998). Measurement reliability and agreement in psychiatry. *Statistical Methods in Medical Research*, 7(3), 301–317. <https://doi.org/10.1177/096228029800700306>
- Silveira-Maia, M., Lopes-Dos-Santos, P., Sanches-Ferreira, M., Alves, S., & Silveira-Maia, C. (2019). The use of the international classification of functioning, disability and health in an interactive perspective: the assessment and intervention of students' additional support needs in Portugal. *Disability and Rehabilitation*, 41(25), 3016–3024. <https://doi.org/10.1080/09638288.2018.1483433>
- Sireci, S., & Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psicothema*, 1(26), 100–107. <https://doi.org/10.7334/psicothema2013.256>
- Sumintono, B., & Widhiarso, W. (2015). Aplikasi pemodelan Rasch pada assessment pendidikan. <http://eprints.um.edu.my/14228/>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Tesio, L. (2003). Measuring Behaviours And Perceptions: Rasch Analysis As a Tool For Rehabilitation Research. *Journal of Rehabilitation Medicine*, 35(3), 105–115. <https://doi.org/10.1080/16501970310010448>
- Werkhoven, S., Anderson, J. H., & Robeyns, I. a. M. (2022). Who benefits from diagnostic labels for developmental disorders? *Developmental Medicine & Child Neurology*, 64(8), 944–949. <https://doi.org/10.1111/dmcn.15177>
- World Health Organization. (2001). *International Classification of Functioning, Disability and Health* [E-Book]. <https://iris.who.int/server/api/core/bitstreams/83e1c3e0-cf9e-4063-a424-71ccc9a84ded/content>
- World Health Organization. (2007). *International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY* [E-Book]. <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health>
- Xavier, S. (2021). Falling through the Cracks: Identifying children with learning difficulties in Malaysian schools. In *IDEAS* (Brief Ideas No. 32). Institute for Democracy and Economic Affaris. <https://www.ideas.org.my/publications-item/brief-ideas-no-32-falling-through-the-cracks-identifying-children-with-learning-difficulties-in-malaysian-schools/>
- Yusoff, M. S. B. (2019). ABC of Content Validation and Content Validity Index Calculation. *Education in Medicine Journal*, 11(2), 49–54. <https://doi.org/10.21315/eimj2019.11.2>

Appendix - I-CVI and κ *Values for All Items**(1a) Body Function – Fairness**

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I-CVI	κ *
1.0 Cognitive Function	1.1	0	1	1	1	1	1	1	1	0.88	0.87
	1.2.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.5.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.5.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.6.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.6.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.2	1	1	1	1	1	1	1	1	1.00	1.00
1.8.3	1	1	1	1	1	1	1	1	1.00	1.00	
1.8.4	1	1	1	1	1	1	1	1	1.00	1.00	
1.8.5	1	1	1	1	1	1	1	1	1.00	1.00	
1.9	1	1	0	1	1	1	1	1	0.88	0.87	
1.10.1	1	1	1	1	1	1	1	1	1.00	1.00	
1.10.2	1	1	1	1	1	1	1	1	1.00	1.00	
2.0 Sensory Function	2.1	1	1	1	1	1	1	1	1	1.00	1.00
	2.2	1	1	1	1	1	1	1	1	1.00	1.00
	2.3	1	1	1	1	1	1	1	1	1.00	1.00
3.0 Language & Speech Function	3.1	1	1	1	1	1	1	1	1	1.00	1.00
	3.2	1	1	1	1	1	1	1	1	1.00	1.00
	3.3	1	1	1	1	1	1	1	1	1.00	1.00
4.0 Cognitive Function	4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.3	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.3	1	1	1	1	1	1	1	1	1.00	1.00
5.0 Other Body Function	5.1	1	1	1	0	1	1	1	1	0.88	0.87
	5.2	1	1	1	0	1	1	1	1	0.88	0.87
	5.3	1	1	0	0	1	1	1	1	0.75	0.72
	5.4	1	1	0	0	1	1	1	1	0.75	0.72
	5.5	1	1	1	1	1	1	1	1	1.00	1.00
	5.6	1	1	1	1	1	1	1	1	1.00	1.00
Proportion relevance	0.98	1	0.93	0.91	1	1	1	1			
Average proportion of items judged as relevance across eight experts										0.98	
S-CVI/Average										0.98	

(1b) Body Function – Relevance

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I-CVI	κ *
1.0 Cognitive Function	1.1	0	1	1	1	1	1	1	0	0.75	0.72
	1.2.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.3	1	1	1	1	0	1	1	1	0.88	0.87
	1.2.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	0	1	1	1	0.88	0.87
	1.4.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.5.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.5.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.6.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.6.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.2	1	1	1	1	1	1	1	1	1.00	1.00
1.8.3	1	1	1	1	1	1	1	1	1.00	1.00	
1.8.4	1	1	1	1	1	1	1	1	1.00	1.00	
1.8.5	1	1	1	1	1	1	1	1	1.00	1.00	
1.9	1	1	0	1	1	1	1	1	0.88	0.87	
1.10.1	1	1	1	1	1	1	1	1	1.00	1.00	
1.10.2	1	1	1	1	1	1	1	1	1.00	1.00	
2.0 Sensory Function	2.1	1	1	1	1	1	1	1	1	1.00	1.00
	2.2	1	1	1	1	1	1	1	1	1.00	1.00
	2.3	1	1	1	1	1	1	1	0	0.88	0.87
3.0 Language & Speech Function	3.1	1	1	1	1	1	1	1	1	1.00	1.00
	3.2	1	1	1	1	1	1	1	1	1.00	1.00
	3.3	1	1	1	1	1	1	1	1	1.00	1.00
4.0 Cognitive Function	4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.3	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.3	1	1	1	1	1	1	1	1	1.00	1.00
5.0 Other Body Function	5.1	1	1	1	0	1	1	1	1	0.88	0.87
	5.2	1	1	1	0	1	1	1	1	0.88	0.87
	5.3	1	1	1	0	1	1	1	1	0.88	0.87
	5.4	1	1	1	0	1	1	1	1	0.88	0.87
	5.5	1	1	1	1	1	1	1	1	1.00	1.00
	5.6	1	1	1	1	1	1	1	1	1.00	1.00
Proportion relevance		0.98	1	0.98	0.91	0.95	1	1	0.95		
Average proportion of items judged as relevance across eight experts										0.97	
S-CVI/Average										0.97	

(1c) Body Function – Clarity

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ *
1.0 Cognitive Function	1.1	1	0	1	1	1	1	1	0	0.75	0.72
	1.2.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.6	1	1	1	1	1	1	1	0	0.88	0.87
	1.2.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	0	1	1	1	0.88	0.87
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.5.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.5.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.6.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.6.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.2	1	1	1	1	1	1	1	1	1.00	1.00
1.8.3	1	1	1	1	1	1	1	1	1.00	1.00	
1.8.4	1	1	1	1	1	1	1	1	1.00	1.00	
1.8.5	1	1	1	1	1	1	1	1	1.00	1.00	
1.9	1	1	1	1	1	1	1	1	1.00	1.00	
1.10.1	1	1	1	1	1	1	1	1	1.00	1.00	
1.10.2	1	1	0	1	1	1	1	1	0.88	0.87	
2.0 Sensory Function	2.1	1	1	1	1	1	1	1	1	1.00	1.00
	2.2	1	1	1	1	1	1	1	1	1.00	1.00
	2.3	1	1	1	1	1	1	1	1	1.00	1.00
3.0 Language and Speech Function	3.1	1	1	1	1	1	1	1	1	1.00	1.00
	3.2	1	1	1	1	1	1	1	1	1.00	1.00
	3.3	1	1	1	1	1	1	1	1	1.00	1.00
4.0 Cognitive Function	4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.3	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.3	1	1	1	1	1	1	1	1	1.00	1.00
5.0 Other Body Function	5.1	1	1	1	0	1	1	1	1	0.88	0.87
	5.2	1	1	1	0	1	1	1	1	0.88	0.87
	5.3	1	1	1	0	1	1	1	1	0.88	0.87
	5.4	1	1	1	0	1	1	1	1	0.88	0.87
	5.5	1	1	1	1	1	1	1	1	1.00	1.00
	5.6	1	1	1	1	1	1	1	1	1.00	1.00
Proportion relevance		1	0.98	0.98	0.91	0.98	1	1	0.95		
Average proportion of items judged as relevance across eight experts										0.97	
S-CVI/Average										0.97	

(1d) Body Function – Ease

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ *
1.0 Cognitive Function	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.2	1	1	0	1	1	1	1	1	0.88	0.87
	1.2.3	1	1	0	1	1	1	1	1	0.88	0.87
	1.2.4	1	1	0	1	1	1	1	1	0.88	0.87
	1.2.5.	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.2.7	1	1	0	1	1	1	1	1	0.88	0.87
	1.3	1	0	0	1	1	1	1	1	0.75	0.72
	1.4.1	1	1	0	1	1	1	1	1	0.88	0.87
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.5.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.5.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.6.1	1	1	0	1	1	1	1	1	0.88	0.87
	1.6.2	1	1	0	1	1	1	1	1	0.88	0.87
	1.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.1	1	1	0	1	1	1	1	1	0.88	0.87
	1.8.2	1	1	1	1	1	1	1	1	1.00	1.00
1.8.3	1	1	0	1	1	1	1	1	0.88	0.87	
1.8.4	1	1	1	1	1	1	1	1	1.00	1.00	
1.8.5	1	1	1	1	1	1	1	1	1.00	1.00	
1.9	1	1	0	1	1	1	1	1	0.88	0.87	
1.10.1	1	1	1	1	1	1	1	1	1.00	1.00	
1.10.2	1	1	1	1	1	1	1	1	1.00	1.00	
2.0 Sensory Function	2.1	1	1	1	1	1	1	1	1	1.00	1.00
	2.2	1	1	1	1	1	1	1	1	1.00	1.00
	2.3	1	1	1	1	1	1	1	1	1.00	1.00
3.0 Language and Speech Function	3.1	1	1	1	1	1	1	1	1	1.00	1.00
	3.2	1	1	1	1	1	1	1	1	1.00	1.00
	3.3	1	1	1	1	1	1	1	1	1.00	1.00
4.0 Cognitive Function	4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.3	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.4.3	1	1	1	1	1	1	1	1	1.00	1.00
5.0 Other Body Function	5.1	1	1	0	0	1	1	1	1	0.75	0.72
	5.2	1	1	0	0	1	1	1	1	0.75	0.72
	5.3	1	1	0	0	1	1	1	1	0.75	0.72
	5.4	1	1	0	0	1	1	1	1	0.75	0.72
	5.5	1	1	1	1	1	1	1	1	1.00	1.00
	5.6	1	1	1	1	1	1	1	1	1.00	1.00
Proportion relevance		1	0.977	0.659	0.909	1	1	1	1		
Average proportion of items judged as relevance across eight experts										0.94	
S-CVI/Average										0.94	

(2a) Activity and Participation – Fairness

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
1.0 Reading Skills	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	0	1	1	1	1	1	1	1	0.88	0.87
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
2.0 Writing Skills	2.1	1	1	1	1	1	1	1	1	1.00	1.00
	2.2	0	1	0	1	1	1	1	1	0.75	0.72
	2.3	0	1	0	1	1	1	1	1	0.75	0.72
	2.4	1	1	1	1	1	1	1	1	1.00	1.00
	2.5	1	1	1	1	1	1	1	1	1.00	1.00
	2.6	1	1	0	1	1	1	1	1	0.88	0.87
3.0 Counting Skills	3.1	1	1	1	1	1	1	1	1	1.00	1.00
	3.2	1	1	1	1	1	1	1	1	1.00	1.00
	3.3	1	1	1	1	1	1	1	1	1.00	1.00
	3.4	1	1	1	1	1	1	1	1	1.00	1.00
	3.5	1	1	1	0	1	1	1	1	0.88	0.87
4.0 Giving Focus and Directing Attention	4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.3	1	1	0	1	1	1	1	1	0.88	0.87
5.0 Thinking Skills	5.1	1	1	1	1	1	1	1	1	1.00	1.00
	5.2	1	1	1	1	1	1	1	1	1.00	1.00
	5.3	1	1	0	0	1	1	1	1	0.75	0.72
6.0 Problem Solving and Making Decision	6.1	1	1	1	1	1	1	1	1	1.00	1.00
	6.2	1	1	0	0	1	1	1	1	0.75	0.72
	6.3	1	1	1	1	1	1	1	1	1.00	1.00
	6.4	1	1	1	1	1	1	1	1	1.00	1.00
7.0 Completing Task	7.1	1	1	1	1	1	1	1	1	1.00	1.00
	7.2	1	1	1	1	1	1	1	1	1.00	1.00
	7.3	1	1	0	1	1	1	1	1	0.88	0.87
	7.4	1	1	0	0	1	1	1	1	0.75	0.72
	7.5	1	1	1	1	1	1	1	1	1.00	1.00
8.0 School Activities	8.1	1	1	1	1	1	1	1	1	1.00	1.00
	8.2	1	1	0	1	1	1	1	1	0.88	0.87
	8.3	1	1	1	1	1	1	1	1	1.00	1.00
	8.4	1	1	1	1	1	1	1	1	1.00	1.00
	8.5	1	1	1	1	1	1	1	1	1.00	1.00
9.0 Communication	9.1	1	1	1	1	1	1	1	1	1.00	1.00
	9.2	1	1	1	1	1	1	1	1	1.00	1.00
	9.3	1	1	0	1	1	1	1	1	0.88	0.87
	9.4	1	1	1	1	1	1	1	1	1.00	1.00
	9.5	1	1	1	1	1	1	1	1	1.00	1.00
	9.6	1	1	0	1	1	1	1	1	0.88	0.87
	9.7	1	1	0	1	1	1	1	1	0.88	0.87
	9.8	1	1	1	1	1	1	1	1	1.00	1.00
	9.9	1	1	1	1	1	1	1	1	1.00	1.00

(2a) continues

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
10.0 Interaction	10.1	1	1	1	1	1	1	1	1	1.00	1.00
	10.2	1	1	1	1	1	1	1	1	1.00	1.00
	10.3	1	1	1	1	1	1	1	1	1.00	1.00
	10.4	1	1	1	1	1	1	1	1	1.00	1.00
	10.5	1	1	1	1	1	1	1	1	1.00	1.00
	10.6	1	1	0	1	1	1	1	1	0.88	0.87
	10.7	1	1	1	1	1	1	1	1	1.00	1.00
	10.8	1	1	1	1	1	1	1	1	1.00	1.00
	10.9	1	1	0	1	1	1	1	1	0.88	0.87
	10.10	1	1	1	1	1	1	1	1	1.00	1.00
11.0 Handling Stress and Other Psychology Needs	11.1	1	1	1	1	1	1	1	1	1.00	1.00
	11.2	1	1	1	1	1	1	1	1	1.00	1.00
	11.3	1	1	1	1	1	1	1	1	1.00	1.00
	11.4	1	1	1	1	1	1	1	1	1.00	1.00
	11.5	1	1	0	1	1	1	1	1	0.88	0.87
	11.6	1	1	0	1	1	1	1	1	0.88	0.87
	11.7	1	1	0	1	1	1	1	1	0.88	0.87
	11.8	1	1	1	1	1	1	1	1	1.00	1.00
	11.9	1	1	1	1	1	1	1	1	1.00	1.00
12.0 Doing Daily Routine	12.1	1	1	1	1	1	1	1	1	1.00	1.00
	12.2	1	1	1	1	1	1	1	1	1.00	1.00
	12.3	1	1	1	1	1	1	1	1	1.00	1.00
	12.4	1	1	1	1	1	1	1	1	1.00	1.00
	12.5	1	1	1	1	1	1	1	1	1.00	1.00
13.0 Body Care	13.1	1	1	0	0	1	1	1	1	0.75	0.72
	13.2	1	1	1	1	1	1	1	1	1.00	1.00
	13.3	1	1	1	1	1	1	1	1	1.00	1.00
	13.4	1	1	0	1	1	1	1	1	0.88	0.87
14.0 Self- Cleaning	14.1	1	1	1	1	1	1	1	1	1.00	1.00
	14.2	1	1	1	1	1	1	1	1	1.00	1.00
	14.3	1	1	1	1	1	1	1	1	1.00	1.00
	14.4	1	1	1	1	1	1	1	1	1.00	1.00
	14.5	1	1	1	1	1	1	1	1	1.00	1.00
	14.6	1	1	1	1	1	1	1	1	1.00	1.00
	14.7	1	1	1	1	1	1	1	1	1.00	1.00
	14.8	1	1	1	1	1	1	1	1	1.00	1.00
15.0 Clothing	15.1	1	1	1	1	1	1	1	1	1.00	1.00
	15.2	1	1	1	1	1	1	1	1	1.00	1.00
	15.3	1	1	1	1	1	1	1	1	1.00	1.00
	15.4	1	1	1	1	1	1	1	1	1.00	1.00
	15.5	1	1	1	1	1	1	1	1	1.00	1.00
	15.6	1	1	0	1	1	1	1	1	0.88	0.87
	15.7	1	1	1	1	1	1	1	1	1.00	1.00
	15.8	1	1	1	1	1	1	1	1	1.00	1.00
16.0 Family Relationship	16.1	1	1	1	1	1	1	1	1	1.00	1.00
	16.2	1	1	1	1	1	1	1	1	1.00	1.00
	16.3	1	1	1	1	1	1	1	1	1.00	1.00

(2a) continues

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
17.0 Recreation and Leisure Time	17.1	1	1	1	1	1	1	1	1	1.00	1.00
	17.2	1	1	1	1	1	1	1	1	1.00	1.00
	17.3	1	1	1	1	1	1	1	1	1.00	1.00
	17.4	1	1	1	1	1	1	1	1	1.00	1.00
	17.5	1	1	1	1	1	1	1	1	1.00	1.00
18.0 Play Activities	18.1	1	1	1	1	1	1	1	1	1.00	1.00
	18.2	1	1	1	1	1	1	1	1	1.00	1.00
	18.3	1	1	1	1	1	1	1	1	1.00	1.00
	18.4	1	1	1	1	1	1	1	1	1.00	1.00
	18.5	1	1	1	1	1	1	1	1	1.00	1.00
19.0 Eat and Drink	19.1	1	1	0	1	1	1	1	1	0.88	0.87
	19.2	1	1	1	1	1	1	1	1	1.00	1.00
	19.3	1	1	1	1	1	1	1	1	1.00	1.00
	19.4	1	1	1	1	1	1	1	1	1.00	1.00
	19.5	1	1	0	1	1	1	1	1	0.88	0.87
	19.6	1	1	0	1	1	1	1	1	0.88	0.87
	19.7	1	1	1	1	1	1	1	1	1.00	1.00
Proportion relevance		0.97	1	0.79	0.95	1	1	1	1		
Average proportion of items judged as relevance across eight experts										0.96	
S-CVI/Average										0.96	

(2b) Activity and Participation – Relevance

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
1.0 Reading Skills	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	0	1	1	1	1	1	1	1	0.88	0.87
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
2.0 Writing Skills	2.1	1	1	1	1	1	1	1	1	1.00	1.00
	2.2	0	1	1	1	1	1	1	1	0.88	0.87
	2.3	0	1	0	1	1	1	1	1	0.75	0.72
	2.4	1	1	1	1	1	1	1	1	1.00	1.00
	2.5	1	1	0	1	1	1	1	1	0.88	0.87
	2.6	1	1	0	1	1	1	1	0	0.75	0.72
3.0 Counting Skills	3.1	1	1	1	1	1	1	1	1	1.00	1.00
	3.2	1	1	1	1	1	1	1	1	1.00	1.00
	3.3	1	1	1	1	1	1	1	1	1.00	1.00
	3.4	1	1	1	1	1	1	1	1	1.00	1.00
	3.5	1	1	1	0	0	1	1	0	0.63	0.68
4.0 Giving Focus and Directing Attention	4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.3	1	1	0	1	1	1	1	1	0.88	0.87
5.0 Thinking Skills	5.1	1	1	1	1	1	1	1	1	1.00	1.00
	5.2	1	1	1	1	1	1	1	1	1.00	1.00
	5.3	1	1	0	0	1	1	1	1	0.75	0.72
6.0 Problem Solving and Making Decision	6.1	1	1	1	1	1	1	1	1	1.00	1.00
	6.2	1	1	0	0	1	1	1	1	0.75	0.72
	6.3	1	1	1	1	1	1	1	1	1.00	1.00
	6.4	1	1	1	1	1	1	1	1	1.00	1.00
7.0 Completing Task	7.1	1	1	1	1	1	1	1	1	1.00	1.00
	7.2	1	1	1	1	1	1	1	1	1.00	1.00
	7.3	1	1	0	1	1	1	1	1	0.88	0.87
	7.4	1	1	0	0	1	1	1	1	0.75	0.72
	7.5	1	1	1	1	1	1	1	1	1.00	1.00
8.0 School Activities	8.1	1	1	1	1	1	1	1	1	1.00	1.00
	8.2	1	1	0	1	1	1	1	1	0.88	0.87
	8.3	1	1	1	1	1	1	1	1	1.00	1.00
	8.4	1	1	1	1	1	1	1	1	1.00	1.00
	8.5	1	1	1	1	1	1	1	1	1.00	1.00
9.0 Communication	9.1	1	1	1	1	1	1	1	1	1.00	1.00
	9.2	1	1	1	1	1	1	1	1	1.00	1.00
	9.3	1	1	0	1	1	1	1	1	0.88	0.87
	9.4	1	1	1	1	1	1	1	1	1.00	1.00
	9.5	1	1	1	1	1	1	1	1	1.00	1.00
	9.6	1	1	0	1	1	1	1	1	0.88	0.87
	9.7	1	1	0	1	1	1	1	1	0.88	0.87
	9.8	1	1	1	1	1	1	1	1	1.00	1.00
	9.9	1	1	1	1	1	1	1	1	1.00	1.00

(2b) continues

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
10.0 Interaction	10.1	1	1	1	1	1	1	1	1	1.00	1.00
	10.2	1	1	1	1	1	1	1	1	1.00	1.00
	10.3	1	1	1	1	1	1	1	1	1.00	1.00
	10.4	1	1	1	1	1	1	1	1	1.00	1.00
	10.5	1	1	1	1	1	1	1	1	1.00	1.00
	10.6	1	1	0	1	1	1	1	1	0.88	0.87
	10.7	1	1	1	1	1	1	1	1	1.00	1.00
	10.8	1	1	1	1	1	1	1	1	1.00	1.00
	10.9	1	1	0	1	1	1	1	1	0.88	0.87
	10.10	1	1	1	1	1	1	1	1	1.00	1.00
11.0 Handling Stress and Other Psychology Needs	11.1	1	1	1	1	1	1	1	1	1.00	1.00
	11.2	1	1	1	1	1	1	1	1	1.00	1.00
	11.3	1	1	1	1	1	1	1	1	1.00	1.00
	11.4	1	1	1	1	1	1	1	1	1.00	1.00
	11.5	1	1	1	1	1	1	1	1	1.00	1.00
	11.6	1	1	1	1	1	1	1	1	1.00	1.00
	11.7	1	1	1	1	1	1	1	1	1.00	1.00
	11.8	1	1	1	1	1	1	1	1	1.00	1.00
	11.9	1	1	1	1	1	1	1	1	1.00	1.00
12.0 Doing Daily Routine	12.1	1	1	1	1	1	1	1	1	1.00	1.00
	12.2	1	1	1	1	1	1	1	1	1.00	1.00
	12.3	1	1	1	1	1	1	1	1	1.00	1.00
	12.4	1	1	1	1	1	1	1	1	1.00	1.00
	12.5	1	1	1	1	1	1	1	1	1.00	1.00
13.0 Body Care	13.1	1	1	0	0	1	1	1	1	0.75	0.72
	13.2	1	1	1	1	1	1	1	1	1.00	1.00
	13.3	1	1	1	1	1	1	1	1	1.00	1.00
	13.4	1	1	0	1	1	1	1	1	0.88	0.87
14.0 Self- Cleaning	14.1	1	1	1	1	1	1	1	1	1.00	1.00
	14.2	1	1	1	1	1	1	1	1	1.00	1.00
	14.3	1	1	1	1	1	1	1	1	1.00	1.00
	14.4	1	1	1	1	1	1	1	1	1.00	1.00
	14.5	1	1	1	1	1	1	1	1	1.00	1.00
	14.6	1	1	1	1	1	1	1	1	1.00	1.00
	14.7	1	1	1	1	1	1	1	1	1.00	1.00
	14.8	1	1	1	1	1	1	1	1	1.00	1.00
15.0 Clothing	15.1	1	1	1	1	1	1	1	1	1.00	1.00
	15.2	1	1	1	1	1	1	1	1	1.00	1.00
	15.3	1	1	1	1	1	1	1	1	1.00	1.00
	15.4	1	1	1	1	1	1	1	1	1.00	1.00
	15.5	1	1	1	1	1	1	1	1	1.00	1.00
	15.6	1	1	0	1	1	1	1	1	0.88	0.87
	15.7	1	1	1	1	1	1	1	1	1.00	1.00
	15.8	1	1	1	1	1	1	1	1	1.00	1.00
16.0 Family Relationship	16.1	1	1	1	1	1	1	1	1	1.00	1.00
	16.2	1	1	1	1	1	1	1	1	1.00	1.00
	16.3	1	1	1	1	1	1	1	1	1.00	1.00

(2b) continues

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
17.0 Recreation and Leisure Time	17.1	1	1	1	1	1	1	1	1	1.00	1.00
	17.2	1	1	1	1	1	1	1	1	1.00	1.00
	17.3	1	1	1	1	1	1	1	1	1.00	1.00
	17.4	1	1	1	1	1	1	1	1	1.00	1.00
	17.5	1	1	1	1	1	1	1	1	1.00	1.00
18.0 Play Activities	18.1	1	1	1	1	1	1	1	1	1.00	1.00
	18.2	1	0	1	1	1	1	1	1	0.88	0.87
	18.3	1	1	1	1	1	1	1	1	1.00	1.00
	18.4	1	1	1	1	1	1	1	1	1.00	1.00
	18.5	1	1	1	1	1	1	1	1	1.00	1.00
19.0 Eat and Drink	19.1	1	1	0	1	1	1	1	1	0.88	0.87
	19.2	1	1	1	1	1	1	1	1	1.00	1.00
	19.3	1	1	1	1	1	1	1	1	1.00	1.00
	19.4	1	1	1	1	1	1	1	1	1.00	1.00
	19.5	1	1	0	1	1	1	1	1	0.88	0.87
	19.6	1	1	0	1	1	1	1	1	0.88	0.87
	19.7	1	1	1	1	1	1	1	1	1.00	1.00
Proportion relevance		0.97	0.99	0.82	0.95	0.99	1	1	0.98		
Average proportion of items judged as relevance across eight experts										0.96	
S-CVI/Average										0.96	

(2c) Activity and Participation – Clarity

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
1.0 Reading Skills	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
2.0 Writing Skills	2.1	1	1	1	1	1	1	1	1	1.00	1.00
	2.2	0	1	1	1	1	1	1	1	0.88	0.87
	2.3	0	1	1	1	1	1	1	1	0.88	0.87
	2.4	1	1	1	1	1	1	1	1	1.00	1.00
	2.5	1	1	1	1	1	1	1	1	1.00	1.00
	2.6	1	1	1	1	1	1	1	1	1.00	1.00
3.0 Counting Skills	3.1	1	1	1	1	1	1	1	1	1.00	1.00
	3.2	1	1	1	1	1	1	1	1	1.00	1.00
	3.3	1	1	1	1	1	1	1	1	1.00	1.00
	3.4	1	1	1	1	1	1	1	1	1.00	1.00
	3.5	1	1	1	1	1	1	1	1	1.00	1.00
4.0 Giving Focus and Directing Attention	4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.3	1	1	1	1	1	1	1	1	1.00	1.00
5.0 Thinking Skills	5.1	1	1	1	1	1	1	1	1	1.00	1.00
	5.2	1	1	1	1	1	1	1	1	1.00	1.00
	5.3	1	1	1	0	1	1	1	0	0.75	0.72
6.0 Problem Solving and Making Decision	6.1	1	1	1	1	1	1	1	1	1.00	1.00
	6.2	1	1	0	0	1	1	1	1	0.75	0.72
	6.3	1	1	1	1	1	1	1	1	1.00	1.00
	6.4	1	1	1	1	1	1	1	1	1.00	1.00
7.0 Completing Task	7.1	1	1	1	1	1	1	1	1	1.00	1.00
	7.2	1	1	1	1	1	1	1	1	1.00	1.00
	7.3	1	1	1	1	1	1	1	1	1.00	1.00
	7.4	1	1	1	0	1	1	1	1	0.88	0.87
	7.5	1	1	1	1	1	1	1	1	1.00	1.00
8.0 School Activities	8.1	1	1	1	1	1	1	1	1	1.00	1.00
	8.2	1	1	0	1	1	1	1	1	0.88	0.87
	8.3	1	1	1	1	1	1	1	1	1.00	1.00
	8.4	1	1	1	1	1	1	1	1	1.00	1.00
	8.5	1	1	1	1	1	1	1	1	1.00	1.00
9.0 Communication	9.1	1	1	1	1	1	1	1	1	1.00	1.00
	9.2	1	1	1	1	1	1	1	1	1.00	1.00
	9.3	1	1	0	1	1	1	1	1	0.88	0.87
	9.4	1	1	1	1	1	1	1	1	1.00	1.00
	9.5	1	1	1	1	1	1	1	1	1.00	1.00
	9.6	1	1	0	1	1	1	1	1	0.88	0.87
	9.7	1	1	0	1	1	1	1	1	0.88	0.87
	9.8	1	1	1	1	1	1	1	1	1.00	1.00
	9.9	1	1	1	1	1	1	1	1	1.00	1.00

(2c) continues

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
10.0 Interaction	10.1	1	1	1	1	1	1	1	1	1.00	1.00
	10.2	1	1	1	1	1	1	1	1	1.00	1.00
	10.3	1	1	1	1	1	1	1	1	1.00	1.00
	10.4	1	1	1	1	1	1	1	1	1.00	1.00
	10.5	1	1	1	1	1	1	1	1	1.00	1.00
	10.6	1	1	0	1	1	1	1	1	0.88	0.87
	10.7	1	1	1	1	1	1	1	1	1.00	1.00
	10.8	1	1	1	1	1	1	1	1	1.00	1.00
	10.9	1	1	0	1	1	1	1	1	0.88	0.87
	10.10	1	1	1	1	1	1	1	1	1.00	1.00
11.0 Handling Stress and Other Psychology Needs	11.1	1	1	1	1	1	1	1	1	1.00	1.00
	11.2	1	1	1	1	1	1	1	1	1.00	1.00
	11.3	1	1	1	1	1	1	1	1	1.00	1.00
	11.4	1	1	1	1	1	1	1	1	1.00	1.00
	11.5	1	1	1	1	1	1	1	1	1.00	1.00
	11.6	1	1	1	1	1	1	1	1	1.00	1.00
	11.7	1	1	1	1	1	1	1	1	1.00	1.00
	11.8	1	1	1	1	1	1	1	1	1.00	1.00
	11.9	1	1	1	1	1	1	1	1	1.00	1.00
12.0 Doing Daily Routine	12.1	1	1	1	1	1	1	1	1	1.00	1.00
	12.2	1	1	1	1	1	1	1	0	0.88	0.87
	12.3	1	0	1	1	1	1	1	1	0.88	0.87
	12.4	1	1	1	1	1	1	1	1	1.00	1.00
	12.5	1	1	1	1	1	1	1	1	1.00	1.00
13.0 Body Care	13.1	1	1	0	0	1	1	1	1	0.75	0.72
	13.2	1	1	1	1	1	1	1	1	1.00	1.00
	13.3	1	1	1	1	1	1	1	1	1.00	1.00
	13.4	1	1	0	1	1	1	1	1	0.88	0.87
14.0 Self- Cleaning	14.1	1	1	1	1	1	1	1	1	1.00	1.00
	14.2	1	1	1	1	1	1	1	1	1.00	1.00
	14.3	1	1	1	1	1	1	1	1	1.00	1.00
	14.4	1	1	1	1	1	1	1	1	1.00	1.00
	14.5	1	1	1	1	1	1	1	1	1.00	1.00
	14.6	1	1	1	1	1	1	1	1	1.00	1.00
	14.7	1	1	1	1	1	1	1	1	1.00	1.00
	14.8	1	1	1	1	1	1	1	1	1.00	1.00
15.0 Clothing	15.1	1	1	1	1	1	1	1	1	1.00	1.00
	15.2	1	1	1	1	1	1	1	1	1.00	1.00
	15.3	1	1	1	1	1	1	1	1	1.00	1.00
	15.4	1	1	1	1	1	1	1	1	1.00	1.00
	15.5	1	1	1	1	1	1	1	1	1.00	1.00
	15.6	1	1	1	1	1	1	1	1	1.00	1.00
	15.7	1	1	1	1	1	1	1	1	1.00	1.00
	15.8	1	1	1	1	1	1	1	1	1.00	1.00
16.0 Family Relationship	16.1	1	1	1	1	1	1	1	1	1.00	1.00
	16.2	1	1	1	1	1	1	1	1	1.00	1.00
	16.3	1	1	1	1	1	1	1	1	1.00	1.00

(2c) continues

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
17.0 Recreation and Leisure Time	17.1	1	1	1	1	1	1	1	1	1.00	1.00
	17.2	1	1	1	1	1	1	1	1	1.00	1.00
	17.3	1	1	1	1	1	1	1	1	1.00	1.00
	17.4	1	1	1	1	1	1	1	1	1.00	1.00
	17.5	1	1	1	1	1	1	1	1	1.00	1.00
18.0 Play Activities	18.1	1	1	1	1	1	1	1	1	1.00	1.00
	18.2	1	0	1	1	1	1	1	1	0.88	0.87
	18.3	1	1	1	1	1	1	1	1	1.00	1.00
	18.4	1	1	1	1	1	1	1	1	1.00	1.00
	18.5	1	1	1	1	1	1	1	1	1.00	1.00
19.0 Eat and Drink	19.1	1	1	0	1	1	1	1	1	0.88	0.87
	19.2	1	1	1	1	1	1	1	1	1.00	1.00
	19.3	1	1	1	1	1	1	1	1	1.00	1.00
	19.4	1	1	1	1	1	1	1	1	1.00	1.00
	19.5	1	1	1	1	1	1	1	1	1.00	1.00
	19.6	1	1	1	1	1	1	1	1	1.00	1.00
	19.7	1	1	1	1	1	1	1	1	1.00	1.00
Proportion relevance		0.98	0.98	0.91	0.96	1	1	1	0.98		
Average proportion of items judged as relevance across eight experts										0.98	
S-CVI/Average										0.98	

(2d) Activity and Participation – Ease

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ *
1.0 Reading Skills	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4	1	1	1	1	0	1	1	1	0.88	0.87
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
2.0 Writing Skills	2.1	1	1	0	1	1	1	1	1	0.88	0.87
	2.2	0	1	0	1	1	1	1	1	0.75	0.72
	2.3	0	1	0	1	1	1	1	1	0.75	0.72
	2.4	1	1	0	1	1	1	1	1	0.88	0.87
	2.5	1	1	0	1	1	1	1	1	0.88	0.87
	2.6	1	1	0	1	1	1	1	1	0.88	0.87
3.0 Counting Skills	3.1	1	1	1	1	1	1	1	1	1.00	1.00
	3.2	1	1	0	1	1	1	1	1	0.88	0.87
	3.3	1	1	0	1	1	1	1	1	0.88	0.87
	3.4	1	1	0	1	1	1	1	1	0.88	0.87
	3.5	1	1	0	0	1	1	1	1	0.75	0.72
4.0 Giving Focus and Directing Attention	4.1	1	1	1	1	1	1	1	1	1.00	1.00
	4.2	1	1	1	1	1	1	1	1	1.00	1.00
	4.3	1	1	1	1	1	1	1	1	1.00	1.00
5.0 Thinking Skills	5.1	1	1	0	1	1	1	1	1	0.88	0.87
	5.2	1	1	1	1	1	1	1	1	1.00	1.00
	5.3	1	1	0	0	1	1	1	0	0.63	0.68
6.0 Problem Solving and Making Decision	6.1	1	1	1	1	1	1	1	1	1.00	1.00
	6.2	1	1	0	0	1	1	1	1	0.75	0.72
	6.3	1	1	1	1	1	1	1	1	1.00	1.00
	6.4	1	1	1	1	1	1	1	1	1.00	1.00
7.0 Completing Task	7.1	1	1	1	1	1	1	1	1	1.00	1.00
	7.2	1	1	1	1	1	1	1	1	1.00	1.00
	7.3	1	1	1	1	1	1	1	1	1.00	1.00
	7.4	1	1	0	0	1	1	1	1	0.75	0.72
	7.5	1	1	0	1	1	1	1	1	0.88	0.87
8.0 School Activities	8.1	1	1	1	1	1	1	1	1	1.00	1.00
	8.2	1	1	0	1	1	1	1	1	0.88	0.87
	8.3	1	1	1	1	1	1	1	1	1.00	1.00
	8.4	1	1	1	1	1	1	1	1	1.00	1.00
	8.5	1	1	1	1	1	1	1	1	1.00	1.00
9.0 Communication	9.1	1	1	1	1	1	1	1	1	1.00	1.00
	9.2	1	1	1	1	1	1	1	1	1.00	1.00
	9.3	1	1	0	1	1	1	1	1	0.88	0.87
	9.4	1	1	1	1	1	1	1	1	1.00	1.00
	9.5	1	1	1	1	1	1	1	1	1.00	1.00
	9.6	1	1	0	1	1	1	1	1	0.88	0.87
	9.7	1	1	0	1	1	1	1	1	0.88	0.87
	9.8	1	1	1	1	1	1	1	1	1.00	1.00
	9.9	1	1	1	1	1	1	1	1	1.00	1.00

(2d) continues

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
10.0 Interaction	10.1	1	1	1	1	1	1	1	1	1.00	1.00
	10.2	1	1	1	1	1	1	1	1	1.00	1.00
	10.3	1	1	0	1	1	1	1	1	0.88	0.87
	10.4	1	1	0	1	1	1	1	1	0.88	0.87
	10.5	1	1	1	1	1	1	1	1	1.00	1.00
	10.6	1	1	0	1	1	1	1	1	0.88	0.87
	10.7	1	1	1	1	1	1	1	1	1.00	1.00
	10.8	1	1	1	1	1	1	1	1	1.00	1.00
	10.9	1	1	0	1	1	1	1	1	0.88	0.87
	10.10	1	1	1	1	1	1	1	1	1.00	1.00
11.0 Handling Stress and Other Psychology Needs	11.1	1	1	1	1	1	1	1	1	1.00	1.00
	11.2	1	1	1	1	1	1	1	1	1.00	1.00
	11.3	1	1	0	1	1	1	1	1	0.88	0.87
	11.4	1	1	1	1	1	1	1	1	1.00	1.00
	11.5	1	1	0	1	1	1	1	1	0.88	0.87
	11.6	1	1	0	1	1	1	1	1	0.88	0.87
	11.7	1	1	0	1	1	1	1	1	0.88	0.87
	11.8	1	1	1	1	1	1	1	1	1.00	1.00
	11.9	1	1	1	1	1	1	1	1	1.00	1.00
12.0 Doing Daily Routine	12.1	1	1	0	1	1	1	1	1	0.88	0.87
	12.2	1	1	1	1	1	1	1	1	1.00	1.00
	12.3	1	1	1	1	1	1	1	1	1.00	1.00
	12.4	1	1	0	1	1	1	1	1	0.88	0.87
	12.5	1	1	0	1	1	1	1	1	0.88	0.87
13.0 Body Care	13.1	1	1	0	0	1	1	1	1	0.75	0.72
	13.2	1	1	1	1	1	1	1	1	1.00	1.00
	13.3	1	1	1	1	1	1	1	1	1.00	1.00
	13.4	1	1	0	1	1	1	1	1	0.88	0.87
14.0 Self- Cleaning	14.1	1	1	1	1	1	1	1	1	1.00	1.00
	14.2	1	1	0	1	1	1	1	1	0.88	0.87
	14.3	1	1	0	1	1	1	1	1	0.88	0.87
	14.4	1	1	0	1	1	1	1	1	0.88	0.87
	14.5	1	1	1	1	1	1	1	1	1.00	1.00
	14.6	1	1	1	1	1	1	1	1	1.00	1.00
	14.7	1	1	0	1	1	1	1	1	0.88	0.87
	14.8	1	1	1	1	1	1	1	1	1.00	1.00
15.0 Clothing	15.1	1	1	1	1	1	1	1	1	1.00	1.00
	15.2	1	1	1	1	1	1	1	1	1.00	1.00
	15.3	1	1	0	1	1	1	1	1	0.88	0.87
	15.4	1	1	0	1	1	1	1	1	0.88	0.87
	15.5	1	1	1	1	1	1	1	1	1.00	1.00
	15.6	1	1	0	1	1	1	1	1	0.88	0.87
	15.7	1	1	1	1	1	1	1	1	1.00	1.00
	15.8	1	1	1	1	1	1	1	1	1.00	1.00
16.0 Family Relationship	16.1	1	1	1	1	1	1	1	1	1.00	1.00
	16.2	1	1	1	1	1	1	1	1	1.00	1.00
	16.3	1	1	1	1	1	1	1	1	1.00	1.00

(2d) continues

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
17.0 Recreation and Leisure Time	17.1	1	1	1	1	1	1	1	1	1.00	1.00
	17.2	1	1	1	1	1	1	1	1	1.00	1.00
	17.3	1	1	1	1	1	1	1	1	1.00	1.00
	17.4	1	1	1	1	1	1	1	1	1.00	1.00
	17.5	1	1	1	1	1	1	1	1	1.00	1.00
18.0 Play Activities	18.1	1	1	1	1	1	1	1	1	1.00	1.00
	18.2	1	1	1	1	1	1	1	1	1.00	1.00
	18.3	1	1	1	1	1	1	1	1	1.00	1.00
	18.4	1	1	1	1	1	1	1	1	1.00	1.00
	18.5	1	1	1	1	1	1	1	1	1.00	1.00
19.0 Eat and Drink	19.1	1	1	0	1	1	1	1	1	0.88	0.87
	19.2	1	1	0	1	1	1	1	1	0.88	0.87
	19.3	1	1	0	1	1	1	1	1	0.88	0.87
	19.4	1	1	0	1	1	1	1	1	0.88	0.87
	19.5	1	1	0	1	1	1	1	1	0.88	0.87
	19.6	1	1	0	1	1	1	1	1	0.88	0.87
	19.7	1	1	1	1	1	1	1	1	1.00	1.00
Proportion relevance		0.98	1	0.59	0.95	0.99	1	1	0.99		
Average proportion of items judged as relevance across eight experts										0.94	
S-CVI/Average										0.94	

(3a) Environmental Factors – Fairness

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
1.0 At School	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.3	1	1	1	1	1	1	1	1	1.00	1.00
1.8	1	1	1	1	1	1	1	1	1.00	1.00	
2.0 At Home	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.7	1	1	1	1	0	1	1	1	0.88	0.87
	1.8.1	1	1	1	1	1	1	0	1	0.88	0.87
	1.8.2	1	1	1	1	1	1	1	1	1.00	1.00
1.9	1	1	1	1	1	1	1	1	1.00	1.00	
1.10	1	1	1	1	1	1	1	1	1.00	1.00	
1.11	1	1	1	1	1	1	1	1	1.00	1.00	
Proportion relevance		1	1	1	1	0.96	1	0.96	1		
Average proportion of items judged as relevance across eight experts										0.99	
S-CVI/Average										0.99	

(3b) Environmental Factors – Relevance

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
1.0 At School	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.3	1	1	1	1	1	1	1	1	1.00	1.00
1.8	1	1	1	1	1	1	1	1	1.00	1.00	
2.0 At Home	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.1	1	1	1	1	1	1	0	1	0.88	0.87
	1.8.2	1	1	1	1	1	1	1	1	1.00	1.00
1.9	1	1	1	1	1	1	1	1	1.00	1.00	
1.10	1	1	1	1	1	1	1	1	1.00	1.00	
1.11	1	1	1	1	1	1	1	1	1.00	1.00	
Proportion relevance		1	1	1	1	1	1	0.96	1		
Average proportion of items judged as relevance across eight experts										0.99	
S-CVI/Average										0.99	

(3c) Environmental Factors – Clarity

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
1.0 At School	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.3	1	1	1	1	1	1	1	1	1.00	1.00
1.8	1	1	1	1	1	1	1	1	1.00	1.00	
2.0 At Home	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.1	1	1	1	1	1	1	0	1	0.88	0.87
	1.8.2	1	1	1	1	1	1	1	1	1.00	1.00
1.9	1	1	1	1	1	1	1	1	1.00	1.00	
1.10	1	1	1	1	1	1	1	1	1.00	1.00	
1.11	1	1	1	1	1	1	1	1	1.00	1.00	
Proportion relevance		1	1	1	1	1	1	0.96	1		
Average proportion of items judged as relevance across eight experts										0.99	
S-CVI/Average										0.99	

(3d) Environmental Factors – Ease

Domain	Item	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	I- CVI	κ^*
1.0 At School	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	0	1	1	1	1	1	0.88	0.87
	1.4.3	1	1	0	1	1	1	1	1	0.88	0.87
	1.5	1	1	0	1	1	1	1	1	0.88	0.87
	1.6	1	1	0	1	1	1	1	1	0.88	0.87
	1.7.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.7.3	1	1	1	1	1	1	1	1	1.00	1.00
1.8	1	1	0	1	1	1	1	1	0.88	0.87	
2.0 At Home	1.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.3	1	1	0	1	1	1	1	1	0.88	0.87
	1.4.1	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.2	1	1	1	1	1	1	1	1	1.00	1.00
	1.4.3	1	1	1	1	1	1	1	1	1.00	1.00
	1.5	1	1	1	1	1	1	1	1	1.00	1.00
	1.6	1	1	1	1	1	1	1	1	1.00	1.00
	1.7	1	1	1	1	1	1	1	1	1.00	1.00
	1.8.1	1	1	0	1	1	1	0	1	0.75	0.72
	1.8.2	1	1	1	1	1	1	1	1	1.00	1.00
1.9	1	1	1	1	1	1	1	1	1.00	1.00	
1.10	1	1	1	1	1	1	1	1	1.00	1.00	
1.11	1	1	1	1	1	1	1	1	1.00	1.00	
Proportion relevance		1	1	0.73	1	1	1	0.96	1		
Average proportion of items judged as relevance across eight experts										0.96	
S-CVI/Average										0.96	