

Advancing Early Neurodevelopmental Screening through a Multidomain Validated Instrument

Suharsiwi¹; Weny Savitry S. Pandia²; Vítor Daniel Ferreira Franco³;
Adiyati Fathu Roshonah⁴; Rohimi Zamzam⁵; Misriandi⁶; Ni'matuzahroh⁷

^{1,4,5,6}Universitas Muhammadiyah Jakarta, Indonesia

²Universitas Katolik Indonesia Atma Jaya, Indonesia

³Universidade de Evora, Portugal

⁷Universitas Muhammadiyah Malang, Indonesia

¹Correspondence Email: suharsiwi@umj.ac.id

Article Info

Received: February 16, 2026
Accepted: September 28, 2026
Published: September 30, 2026
Online First: September 30, 2026

Keywords

Developmental Screening;
Inclusive Education;
Multidomain Assessment;
Neurodevelopmental Difficulties;
Early Identification.

Abstract

Early identification of neurodevelopmental difficulties is essential for timely educational support, yet existing screening approaches often assess developmental domains separately and provide limited contextual information for educational decision-making. This study developed and conducted preliminary validation of a multidomain developmental screening instrument for use in educational settings. Using a research and development design, the study involved a needs assessment with 103 participants comprising 65 elementary school teachers and 38 parents, followed by instrument development and content validation by seven multidisciplinary experts. The resulting framework integrated contextual information with 99 screening items in five developmental domains: cognitive development, behavioural regulation, sensory processing, communication, and adaptive functioning. The needs assessment indicated a strong demand for comprehensive developmental screening, while the expert evaluation yielded an overall content-validation score of 84.64%. Expert judgments also demonstrated strong consistency, with an average-measure Intraclass Correlation Coefficient of 0.949. These findings provided preliminary evidence supporting the instrument's content and multidomain structure. The instrument therefore offers a context-sensitive framework for early developmental screening in educational settings, while further field-based psychometric validation is required before broader implementation.

A. Introduction

Failure to identify neurodevelopmental difficulties during early childhood can delay access to educational and developmental support at a period when intervention is particularly consequential. Neurodevelopmental difficulties have become one of the most significant global concerns at the intersection of child development, education, and public health. These conditions encompass a broad range of developmental challenges affecting cognitive functioning, communication, social interaction, behavioural regulation, sensory processing, adaptive behaviour, attention, and learning abilities (Boulton et al., 2023; Colizzi et al., 2020). International evidence indicates that early identification and intervention substantially improve developmental outcomes, educational participation, and long-term quality of life for children with neurodevelopmental difficulties (Kaiser et al., 2024; Warren et al., 2016). Consequently, early developmental screening has become a fundamental component of inclusive educational and child health services worldwide. Responding effectively therefore requires an integrated perspective combining developmental psychology, inclusive education, child health, and educational assessment.

The importance of early developmental screening is closely associated with the rapid neuroplasticity that characterizes early childhood. This critical developmental period provides a valuable opportunity for early identification, enabling timely educational intervention and individualized support that can positively influence children's developmental trajectories (Darling-Hammond et al., 2020). From a developmental psychology perspective, neuroplasticity establishes the biological rationale for why early detection matters; from an inclusive education perspective, it establishes the pedagogical imperative to identify needs before they become entrenched barriers to participation; and from a child health perspective, it defines the window within which preventive and remedial services yield the greatest return (Britto et al., 2017; Ediyanto et al., 2021; Lipkin & Macias, 2020). Developmental screening is not intended to provide a clinical diagnosis; rather, it serves as an initial process for identifying children who may require further comprehensive assessment (Choo et al., 2019).

Despite this established importance, a substantial gap remains between the principle of early identification and its implementation in educational settings. In many schools, early identification still relies heavily on informal teacher observations and non-standardized assessments, resulting in considerable variability in assessment outcomes (Lipkin & Macias, 2020; Schonhaut et al., 2021). Limited access to developmental

specialists, insufficient screening resources, and the absence of culturally appropriate assessment instruments further complicates the identification process. In Indonesia, these challenges are intensified by the diversity of developmental needs within inclusive educational settings and the limited availability of validated screening instruments designed for school-based use. These implementation realities reveal a critical disconnect: while developmental psychology and child health research have advanced considerably in understanding the multidimensional nature of neurodevelopmental difficulties, educational assessment practices in inclusive settings have not kept pace, leaving teachers without empirically grounded tools that reflect this multidimensionality.

Several developmental screening instruments have been widely used internationally, including the Denver Developmental Screening Test, Ages and Stages Questionnaires, Parents' Evaluation of Developmental Status, Modified Checklist for Autism in Toddlers, and Strengths and Difficulties Questionnaire. These instruments have demonstrated satisfactory psychometric properties and have been successfully validated across different populations and cultural contexts, supporting their utility for early developmental identification (Shum et al., 2022; Velikonja et al., 2017). However, previous studies also reveal several limitations. Many existing instruments focus primarily on specific developmental domains, such as autism symptoms, behavioral problems, language development, or socio-emotional functioning, rather than providing a comprehensive multidimensional assessment (Khandare & Goyal, 2024). In addition, several screening tools were developed within Western sociocultural contexts and may require substantial adaptation before implementation in different educational environments.

These limitations, however, point to a deeper conceptual and practical problem that extends beyond the mere absence of a comprehensive instrument. The prevailing domain-specific orientation of existing screening tools can result in a fragmented understanding of neurodevelopmental difficulties, whereby cognitive, communicative, sensory, behavioral, social, and adaptive functioning are assessed in isolation rather than as interrelated dimensions of a child's developmental profile. This creates an important conceptual gap: screening conducted without an integrated framework may provide an incomplete account of developmental risk and its educational implications. Practically, even when individual domains are screened adequately, the absence of contextual information such as developmental history, family background, and health

status restricts the utility of screening results for educational decision-making within inclusive classrooms.

More specifically, the research gap concerns the limited integration of multidomain developmental assessment, contextual child information, and school-based educational use within a single screening framework. Relatively few studies have examined how a multidomain integrated approach can support the quality and educational relevance of early identification, and fewer still have systematically developed and expert-validated such instruments for use by teachers in inclusive educational settings (Boateng et al., 2018; Noordink et al., 2025; DeVellis, 2020). The existing body of knowledge, therefore, provides limited evidence on how integrated multidomain screening can operationally connect developmental science with classroom-based identification practices. This gap is particularly consequential in educational contexts where specialist assessment resources are limited, and teachers frequently constitute the first point of developmental observation.

The present study addresses these gaps by developing and validating an Early Screening Instrument for Neurodevelopmental Difficulties in Children. Unlike previous screening instruments that primarily assess isolated developmental domains, the proposed instrument integrates developmental history, family and health information, cognitive abilities, communication skills, sensory processing, behavioral characteristics, social interaction, and adaptive functioning into a unified assessment framework. The novelty of this study lies not only in the technical construction of a multidomain screening tool but, more fundamentally, in its conceptual contribution: it operationalizes the principle that neurodevelopmental difficulties are best understood and identified through the simultaneous consideration of multiple interacting developmental dimensions within the child's ecological context. Accordingly, screening is positioned as an integrated and context-sensitive educational process rather than solely as a domain-specific detection procedure (Lipkin & Macias, 2020; Schonhaut et al., 2021).

In terms of its contributions, this study offers both theoretical and practical advancements to the field. Theoretically, it provides a comprehensive conceptual framework for developmental screening that reflects multidimensional and ecologically grounded understandings of neurodevelopmental difficulties, thereby extending the conceptual basis of school-based early identification beyond isolated developmental domains. It also contributes preliminary empirical evidence regarding content validity and agreement among expert raters through systematic expert evaluation, addressing a

methodological gap in the developmental screening literature. Practically, the study provides a structured assessment instrument intended to support teachers in conducting early identification within inclusive classrooms while facilitating collaboration among educators, families, psychologists, and healthcare professionals. In this way, the study connects developmental knowledge with educational assessment and inclusive classroom practice without positioning the instrument as a substitute for comprehensive clinical diagnosis.

This study aims to develop and validate an Early Screening Instrument for Neurodevelopmental Difficulties in Children. Specifically, the study seeks to construct comprehensive developmental screening items, evaluate their content validity through expert review, and examine inter-rater reliability among expert judgments as an initial stage of instrument validation. By doing so, the study establishes a preliminary empirical foundation for a multidomain, context-sensitive screening framework intended for subsequent field validation and implementation within inclusive educational settings.

B. Method

This study employed a Research and Development (R&D) design adapted from Borg and Gall (2007), focusing on the initial stages of instrument development: needs assessment, item construction, expert validation, revision, and preliminary reliability evaluation (Gall et al., 2007). The study focused on establishing preliminary evidence of content validity and consistency of expert judgments rather than evaluating intervention effectiveness or the instrument's full psychometric properties. Central to the design was an interdisciplinary approach integrating developmental psychology, special education, early childhood education, and inclusive education, ensuring that the selected constructs and indicators reflected the multidimensional nature of neurodevelopmental difficulties.

Two groups of participants were selected through purposive sampling. The first comprised 103 respondents (65 elementary school teachers and 38 parents) who participated in the needs assessment. Teachers were recruited from inclusive and regular elementary schools across various regions of Indonesia, while parents were recruited from families with school-aged children. They were selected because of their direct involvement in children's educational and developmental processes. The second group consisted of seven experts representing psychology, special education, early childhood education, and inclusive education. All experts held at least a master's degree

and had a minimum of five years of professional experience in child development, developmental assessment, or inclusive education practice.

The instrument development followed contemporary guidelines for systematic instrument construction (Boateng et al., 2018; Kyriazos & Stalikas, 2018; Slavec & Drnovšek, 2012). A structured questionnaire was administered to teachers and parents to identify current developmental screening practices, challenges in early identification, and expectations regarding the content and format of a comprehensive screening tool. Descriptive statistics were used to summarize the findings, which were then integrated with literature on child development, neurodevelopmental difficulties, developmental screening, and inclusive education. Operational indicators were formulated as observable behavioral and functional markers and converted into screening items covering background information and five core domains: cognitive development, behavioral regulation, sensory processing, communication, and adaptive functioning.

During expert validation, the seven experts evaluated each screening item using a structured form addressing content relevance, clarity, representativeness of developmental constructs, and practical applicability (Roebianto et al., 2023; Romero Jeldres et al., 2023). Experts provided quantitative ratings and qualitative comments. Items with low relevance ratings, ambiguous wording, redundancy, or inadequate domain representation were revised accordingly. The revised item set was subsequently evaluated for agreement among expert raters. Content validity was determined through percentage agreement across expert ratings. Inter-rater reliability of expert judgments was evaluated using the Intraclass Correlation Coefficient (ICC), employing a two-way random-effects model with absolute agreement based on mean ratings (Hair et al., 2019; Koo & Li, 2016). Thus, the ICC was interpreted as evidence of consistency among expert judgments rather than as evidence of the instrument's reliability in its intended population.

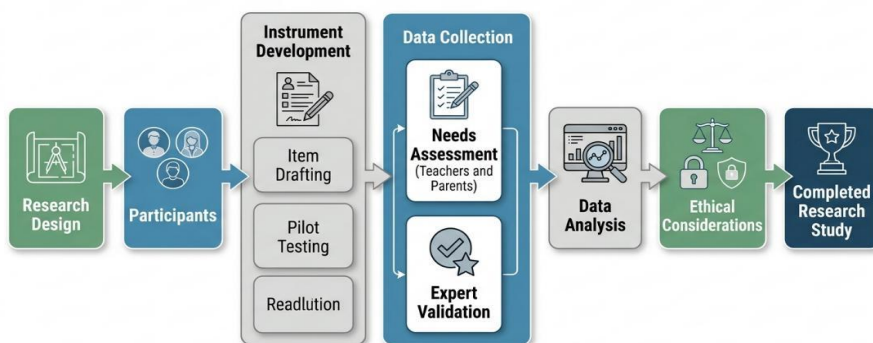


Figure 1. Research Design and Methodology Flowchart

All participants were informed about the study's objectives and procedures before participation. Participation was voluntary, and informed consent was obtained from all participants. Confidentiality and anonymity were maintained throughout the research process, and all collected data were used exclusively for academic purposes.

C. Results and Discussion

This section presents the findings derived from the systematic research and development process. It begins with the needs assessment involving teachers and parents, followed by the conceptual development and structure of the screening instrument. Subsequently, it details the expert validation outcomes, substantive item revisions, and the final reliability testing phase.

1. Results

The results are presented according to the sequential stages of instrument development described in the Method section. They comprise the needs assessment involving teachers and parents, development of the multidomain screening instrument, expert content validation and item revision, assessment of agreement among expert raters, and the resulting proposed developmental screening process.

a. Needs Assessment

A needs assessment was conducted involving 103 respondents (65 teachers, 63.1%; 38 parents, 36.9%) to identify the necessity for a comprehensive developmental screening instrument for the early identification of neurodevelopmental difficulties in children. The questionnaire explored respondents' perceptions regarding the need for structured developmental screening, comprehensive developmental domains, usability, implementation readiness, and the expected benefits of the proposed instrument.

Table 1. Characteristics of Respondents

Variable	Category	n	%
Respondent	Teacher	65	63.1
	Parent	38	36.9
Education	High School	71	68.9
	Bachelor's Degree	32	31.1

The respondents represented diverse geographical regions across Indonesia, including DKI Jakarta, West Java, West Kalimantan, Kepulauan Bangka Belitung, North Sumatra, East Kalimantan, Central Sulawesi, Southeast Sulawesi, Aceh, and East Java. Thus, the needs assessment incorporated perspectives from teachers and parents across multiple regional contexts rather than from a single geographical setting.

Table 2. Needs Assessment Results

Dimension	Items	Mean $\pm$ SD	Score (%)	Category
Need for Developmental Screening	8	4.36 $\pm$ 0.52	87.2	Very High
Comprehensive Developmental Domains	7	4.28 $\pm$ 0.48	85.6	Very High
Usability	6	4.18 $\pm$ 0.57	83.6	High
Implementation Readiness	5	4.10 $\pm$ 0.61	82.0	High
Perceived Benefits	8	4.42 $\pm$ 0.46	88.4	Very High
Overall	34	4.27 $\pm$ 0.47	85.4	Very High

Note. Score (%) represents the mean score relative to the maximum Likert-scale score of 5, calculated as $(\text{Mean}/5) \times 100$. Agreement percentages reported in the text represent the proportion of respondents selecting “Agree” or “Strongly Agree.”

The overall needs assessment score indicated a high level of demand for the development of a comprehensive developmental screening instrument. Among the five assessed dimensions, Perceived Benefits obtained the highest mean score ($M = 4.42$, $SD = 0.46$), with 93.1% of respondents expressing agreement or strong agreement that a comprehensive instrument would improve early identification practices. This was followed by Need for Developmental Screening ($M = 4.36$, $SD = 0.52$; 91.3% agreement). The Comprehensive Developmental Domains dimension also received substantial support ($M = 4.28$, $SD = 0.48$; 88.7% agreement), with 86.4% of respondents specifically endorsing the inclusion of multiple developmental domains.

Regarding Implementation Readiness, 79.6% of respondents indicated willingness to use a structured screening instrument, although 67.3% noted the need for training and simplified administration procedures. Taken together, the needs-assessment findings identified three practical requirements that informed subsequent instrument development: multidomain developmental coverage, usability in educational contexts, and sufficiently clear administration procedures. These findings served as empirical input for instrument construction rather than as evidence of the instrument’s effectiveness.

b. Development of the Screening Instrument

The developmental screening instrument was developed through the integration of three complementary sources: (1) a comprehensive review of contemporary literature on child development, neurodevelopmental difficulties, and developmental screening; (2) findings from the needs assessment involving teachers and parents; and (3) expert consultation throughout the instrument development process. These sources informed the initial instrument structure, operational indicators, and observable screening items subsequently subjected to expert evaluation.

The resulting instrument was organized into two principal sections. The first section comprises background information, including the child's profile, parent or guardian profile, socioeconomic characteristics, child health history, developmental history, home habits, and family health history. These constitute seven contextual or background-information components. The second section comprises 99 screening items distributed across five core developmental domains: cognitive development, behavioral regulation, sensory processing, communication, and adaptive functioning.

The number of items varied across the five developmental domains according to the breadth and number of observable indicators represented within each construct. Sensory processing contains 30 items covering visual, auditory, tactile, olfactory, gustatory, vestibular, and proprioceptive responses. Cognitive development contains 24 items representing attention, memory, problem-solving, concept formation, and time concepts, while behavioral regulation contains 24 items addressing rigidity, repetitive behaviors, aggression, stereotypic behaviors, and emotional regulation. Communication contains seven items representing expressive and receptive communication, whereas adaptive functioning contains 14 items addressing self-care and daily living skills.

Table 3. Structure of the Developmental Screening Instrument

Domain	Purpose of Assessment	Number of Items	Empirical/ Conceptual Basis
Child Profile	Demographic and educational information	7	Contextual baseline information
Parent/Guardian Profile	Family and socioeconomic characteristics	11	Family context and developmental environment
Socioeconomic Profile	Family socioeconomic conditions	2	Risk factor identification

Domain	Purpose of Assessment	Number of Items	Empirical/ Conceptual Basis
Child Health History	Prenatal, perinatal, and medical history	9	Biological risk factors
Developmental History	Developmental milestones and daily developmental patterns	6	Developmental trajectory documentation
Home Habits	Daily routines and home environment	4	Environmental and behavioral context
Family Health History	Familial developmental and medical risk factors	3	Hereditary and familial risk indicators
Cognitive Development	Attention, memory, problem-solving, concept formation, and time concepts	24	Five represented sub-constructs
Behavioural Regulation	Rigidity, repetitive behaviours, aggression, stereotypic behaviours, and emotional regulation	24	Five represented sub-constructs
Sensory Processing	Visual, auditory, tactile, olfactory, gustatory, vestibular, and proprioceptive responses	30	Seven sensory systems requiring multiple indicators
Communication	Expressive and receptive communication characteristics	7	Two sub-domains with fewer distinguishable indicators
Adaptive Functioning	Self-care and daily living skills	14	Two sub-domains (self-care and daily living)

As detailed in Table 3, the resulting instrument combines 42 background-information items across seven contextual components and 99 screening items distributed across five developmental domains. The background-information section was designed to capture relevant contextual characteristics surrounding the child, family, and developmental environment, whereas the screening section focuses specifically on observable indicators representing the targeted developmental constructs. This structural separation enables contextual information to be documented independently from the developmental indicators used in the screening process. Overall, the organization of these two sections provides a systematic framework for the instrument and constitutes the principal output of the instrument-development stage prior to expert validation and subsequent refinement.

Following the identification of the conceptual domains, a comprehensive framework was established to illustrate the structure and organization of the developmental

screening instrument. Figure 2 summarizes how the needs assessment, literature-informed construct development, and expert input contributed to the resulting instrument structure and distinguishes contextual information from the five core developmental domains.

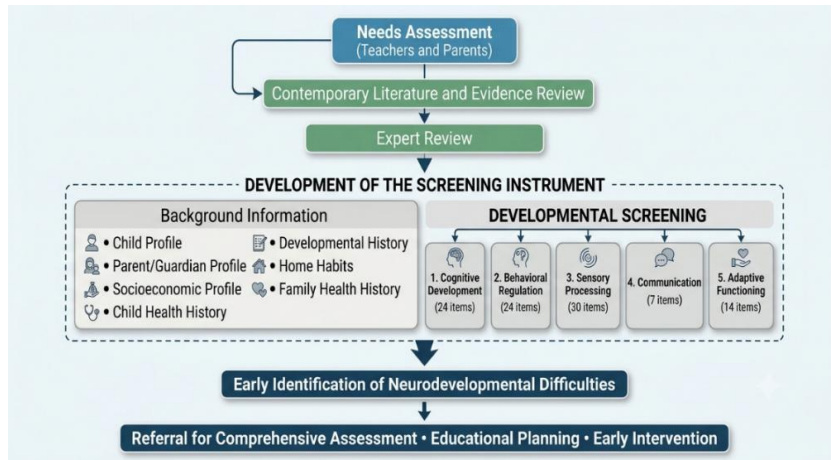


Figure 2. Conceptual Framework for the Development of the Screening Instrument

The framework illustrates the integration of evidence from the needs assessment, contemporary literature, and expert validation in developing a multidimensional developmental screening instrument. The instrument comprises background information and five core developmental domains—cognitive development, behavioral regulation, sensory processing, communication, and adaptive functioning—which formed the substantive structure subsequently subjected to expert evaluation.

c. Expert Validation of the Screening Instrument

The preliminary version of the developmental screening instrument was evaluated by seven experts with expertise in psychology, special education, early childhood education, and inclusive education. The experts assessed the relevance, comprehensiveness, clarity, and representativeness of the instrument across eight assessment domains, including child information, family information, health history, cognitive development, behavioral regulation, sensory processing, communication, and adaptive functioning.

Table 4. Expert Validation Results of the Developmental Screening Instrument

Assessment Domain	Percentage (%)	Category
Child Information	82.86	Very Valid
Family Information	85.71	Very Valid

Assessment Domain	Percentage (%)	Category
Health History	85.71	Very Valid
Cognitive Development	77.14	Valid
Behavioral Regulation	88.57	Very Valid
Sensory Processing	88.57	Very Valid
Communication	85.71	Very Valid
Adaptive Functioning	82.86	Very Valid
Overall	84.64	Very Valid

The expert evaluation produced an overall content-validation score of 84.64%. Behavioral regulation and sensory processing obtained the highest scores (88.57%), followed by family information, health history, and communication (85.71% each). Child information and adaptive functioning each obtained 82.86%, whereas cognitive development received the lowest score (77.14%). The domain-level scores therefore ranged from 77.14% to 88.57%, demonstrating variation in expert evaluations across components of the instrument. The comparatively lower score for cognitive development also provided a specific basis for further revision of items within this domain.

Table 5. Summary of Expert Recommendations for Instrument Revision

Domain	Expert Recommendation	Substantive Revision Implemented
Cognitive Development	Split complex items into simpler observable behaviours	Compound items ("identifies colours and numbers") were separated into individual items; 6 items split; 2 items added.
Attention	Increase attention duration criterion	Duration threshold revised from 3 minutes to 5 minutes; wording clarified to specify task type.
Concept Development	Separate number, colour, and letter recognition	Three compound items were divided into seven independent items with specific observable criteria.
Time Concept	Separate understanding of temporal concepts from sequencing activities	Two items restructured: one assessing temporal vocabulary and one assessing activity sequencing.
Adaptive Functioning	Divide toileting, dressing, and eating skills into independent items	Three multi-component items separated into eight independent items targeting single self-care behaviours.

The experts suggested several substantive revisions to improve the clarity, specificity, and observability of the screening items. Within cognitive development, six compound items were separated and two additional items were introduced. Three compound concept-development items were reorganized into seven independent items, while two time-concept items were restructured to distinguish temporal vocabulary from

activity sequencing. Within adaptive functioning, three multi-component items addressing toileting, dressing, and eating were separated into eight independent items focusing on individual self-care behaviours.

Additionally, experts recommended that ambiguous terms such as “often” or “sometimes” be replaced with frequency anchors (e.g., “more than three times per week”) to reduce subjective interpretation. The attention-duration criterion was revised from three to five minutes, and its wording was clarified to specify the type of task being observed. Thus, expert evaluation resulted not only in linguistic refinement but also in changes to item structure, behavioral specificity, frequency criteria, and observability. All flagged items were substantively revised, and the revised instrument was subsequently evaluated for agreement among expert raters.

Table 6. Reliability Analysis of the Developmental Screening Instrument

Reliability Indicator	Value	Interpretation
Single Measure ICC	0.700	Good
Average Measure ICC	0.949	Excellent

Note: ICC was calculated using a two-way random-effects model with absolute agreement based on mean ratings.

Inter-rater reliability was evaluated using the Intraclass Correlation Coefficient (ICC) based on a two-way random-effects model with absolute agreement (Koo & Li, 2016). The single-measure ICC was 0.700 (95% CI [0.582, 0.814]), indicating good agreement for individual expert ratings. The average-measure ICC was 0.949 (95% CI [0.912, 0.973]), indicating excellent agreement when ratings from all seven experts were averaged. The difference between the two coefficients indicates stronger consistency for aggregated expert judgments than for ratings obtained from an individual expert.

These ICC results provide evidence of consistency among expert judgments during the content-validation stage and complement the domain-level validation results reported above. At this stage, however, the coefficients represent inter-rater agreement among experts rather than reliability of scores obtained from the instrument’s intended users. Accordingly, internal consistency, test-retest reliability, construct validity, criterion validity, and screening accuracy remain outside the scope of the present results and require subsequent field-based psychometric evaluation.

d. Proposed Developmental Screening Process

To illustrate the practical application of the proposed instrument, Figure 3 presents the developmental screening process. Following needs assessment, instrument

construction, expert validation, item revision, and assessment of expert agreement, the resulting instrument was organized into a proposed screening workflow. The workflow integrates information from teachers and parents with the contextual and developmental components established during instrument development.

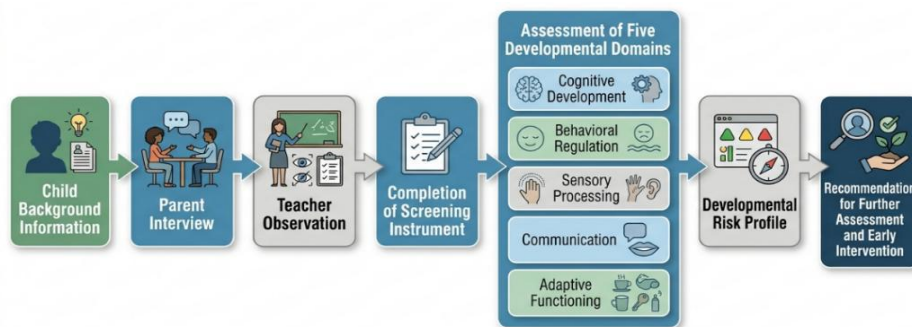


Figure 3. Developmental Screening Process Using the Proposed Instrument

As illustrated in Figure 3, the proposed screening process integrates information from multiple informants, including teachers and parents, to provide a comprehensive profile of children's developmental functioning. Rather than serving as a diagnostic procedure, the instrument is positioned at this stage as a content-validated preliminary screening framework for identifying developmental concerns that may warrant further assessment. The proposed workflow therefore represents the practical organization of the instrument resulting from the present development and expert-validation stages, rather than a fully validated diagnostic or clinical decision-making protocol. Decisions concerning educational or clinical intervention should be supported by appropriate follow-up assessment.

2. Discussion

The present study aimed to develop and validate a contemporary, multidomain developmental screening instrument for the early identification of neurodevelopmental difficulties in children. The findings reveal a coherent pattern across the successive stages of development: teachers and parents expressed substantial demand for structured and multidomain screening; the integration of needs assessment, literature, and expert input produced a framework combining contextual information with five developmental domains; expert review led to substantive refinement of the items; and the resulting expert judgments demonstrated strong consistency. The overall content-

validation score of 84.64% and average-measure ICC of 0.949 therefore provide preliminary support for the content and structure of the proposed instrument. However, these findings represent an early stage of validation rather than evidence of complete psychometric adequacy. Instrument-development literature similarly emphasizes that content validation establishes an important foundation that must subsequently be tested through broader empirical procedures (Boateng et al., 2018; Kyriazos & Stalikas, 2018).

What makes these findings particularly meaningful is that the demand identified by teachers and parents was not limited to the availability of another screening instrument. Respondents simultaneously emphasized comprehensive developmental coverage, usability, implementation readiness, and practical benefits. This suggests that the central problem is not simply whether developmental difficulties can be identified, but whether developmental information can be organized in a form that is sufficiently comprehensive and usable within educational practice. Developmental surveillance requires more than recognizing isolated signs; Lipkin and Macias (2020) emphasize connecting developmental concerns with systematic screening and appropriate follow-up. Recent Indonesian evidence similarly shows that early identification of learning difficulties remains closely connected to teachers' capacity to recognize diverse learner needs and coordinate appropriate support (Yunitasari et al., 2025). The present findings therefore support a screening orientation that connects developmental observation with structured pathways for further assessment rather than treating screening as an endpoint.

This orientation is important because children's developmental functioning rarely fits neatly within isolated categories of observation. The proposed instrument brings cognitive development, behavioral regulation, sensory processing, communication, and adaptive functioning into the same structure while incorporating developmental history, health, family, socioeconomic, and home-related information. Its contribution therefore lies less in the number of domains than in the relationship established between developmental indicators and contextual information. Darling-Hammond et al., (2020) argue that learning and development emerge through dynamic interactions among individual capacities, relationships, experiences, and environments. From a clinical-developmental perspective, Lord et al., (2020) likewise demonstrates the heterogeneity with which neurodevelopmental characteristics can be expressed. This contextual orientation is also compatible with evidence that the educational experiences of students

with disabilities are shaped by interactions among individual characteristics, family background, school structures, and social environments (Supriyanto et al., 2025).

Such integration changes how screening information can be interpreted educationally. A difficulty observed in behavioral regulation may coexist with sensory, communication, cognitive, or contextual factors, while an adaptive difficulty may acquire different meanings depending on developmental history and environmental demands. The present study does not establish causal relationships among these domains, and such relationships should not be inferred from the current data. Nevertheless, organizing them within one framework provides a basis for examining cross-domain patterns. Zwaigenbaum et al., (2015) emphasize that early identification benefits from attention to multiple developmental indicators rather than dependence on a single sign. The multidomain structure consequently functions as a conceptual bridge between observable educational behavior and the broader developmental context in which that behavior occurs, while preserving the distinction between screening and diagnosis.

The expert-validation process further strengthens this interpretation because it did more than generate a numerical judgment of content adequacy. Expert recommendations changed the operational form of several constructs by separating compound behaviors, clarifying attention criteria, replacing ambiguous frequency expressions, and distinguishing individual self-care behaviors. These modifications are methodologically significant because screening items must translate complex developmental constructs into observations that can be interpreted consistently. Boateng et al., (2018) identify item generation and content validity as foundational stages in developing defensible measurement instruments, while Roebianto et al., (2023) emphasize whether items adequately represent their intended content. The expert-review process therefore contributed to the substantive definition of what should be observed, not merely to linguistic refinement. This is particularly important when developmental concepts are translated into instruments intended for educational rather than specialist clinical settings.

The variation in expert-validation scores also shows why the overall score should not be interpreted as evidence that every domain has reached equivalent measurement maturity. Behavioral regulation and sensory processing received the highest evaluations, whereas cognitive development received the lowest. The latter

encompasses attention, memory, problem-solving, concept formation, and temporal understanding, making its translation into discrete observable indicators particularly complex. The restructuring of cognitive items should therefore be understood as construct refinement rather than evidence that the construct has already been psychometrically established. Romero Jeldres et al., (2023) similarly stress that content validation requires attention to specific content representation rather than dependence on an aggregate index alone. This domain-level variation is scientifically useful because it identifies where subsequent empirical testing requires particular attention.

The findings consequently support an ecological and multi-domain perspective on developmental screening without requiring the stronger claim that the study has established a new developmental theory. The conceptual contribution lies instead in operationalizing connections among developmental psychology, child health, educational assessment, and inclusive education within a school-oriented framework. Developmental outcomes are shaped by interconnected psychological, relational, and environmental processes (Darling-Hammond et al., 2020), while developmental conditions can exhibit considerable heterogeneity (Lord et al., 2020). Bringing contextual information into the same framework as developmental indicators therefore allows early identification to be conceptualized as a context-sensitive educational process. This resonates with recent inclusive-education research in Indonesia showing that meaningful inclusion requires educational arrangements responsive to learners' distinctive characteristics rather than standardized treatment of learner diversity (Abdillah et al., 2025).

Within this framework, teachers should not be repositioned as substitute clinicians. Their contribution is more appropriately understood as systematic observation within children's everyday learning environments. Parents provide a complementary perspective because developmental behaviors may differ across home and school contexts. Integrating these sources can make developmental concerns more explicit and communicable when professional referral is required. Ediyanto et al., (2021) show that teachers occupy an important position in implementing inclusive education in Indonesia. Nurdin et al., (2024) likewise found that teachers' limited knowledge in identifying students with special needs remains an obstacle to developing and implementing inclusive educational provision. Developmental surveillance also depends on continuing observation and communication rather than a single assessment event (Lipkin & Macias, 2020). Screening can therefore support

recognition and referral, whereas diagnostic conclusions remain the responsibility of appropriately qualified professionals.

This distinction becomes especially important when the findings are considered beyond Indonesia. Delayed identification of neurodevelopmental conditions is not exclusively an Indonesian problem. Boulton et al., (2023) documented diagnostic delays among children attending a publicly funded developmental assessment service in Australia, demonstrating that difficulties in timely identification can persist even within comparatively well-resourced systems. Kaiser et al., (2024) similarly highlight continuing challenges in connecting early language screening with improved developmental outcomes. The broader international issue is therefore not simply the absence of screening instruments, but the difficulty of constructing effective pathways from early observation to meaningful developmental support. The present study contributes to this wider problem by bringing multidomain developmental information, contextual factors, and multiple informants together at the initial screening stage.

Its international relevance should nevertheless not be framed as the direct export of an Indonesian instrument to other countries. Rather, the potentially transferable contribution lies in the development logic: beginning with locally identified needs, integrating multidisciplinary knowledge, incorporating contextual information, subjecting items to expert scrutiny, and distinguishing preliminary screening from professional diagnosis. Shum et al., (2022), for instance, demonstrated the importance of validating developmental screening within the linguistic and cultural context of Chinese preschoolers in Hong Kong. Evidence from contemporary cross-cultural educational research similarly emphasizes that educational practices cannot simply be transferred across settings without attention to linguistic, cultural, and contextual differences (Mohiyeddini, 2024). Developmental constructs may therefore have broad relevance, but the indicators through which they are recognized and interpreted cannot automatically be assumed to function identically across populations.

This point is particularly important for countries where educational systems must respond to developmental diversity while specialist resources remain unevenly distributed. The framework developed here suggests that contextualized screening may provide a structured interface between what teachers and parents observe and what specialists subsequently need to investigate. Recent Indonesian research demonstrates that schools implementing inclusive education continue to encounter institutional readiness and support challenges, including limitations in programs

designed for children with special needs (Apriliansi et al., 2024). At the classroom level, Ratnawulan et al., (2025) similarly show that educating children with special needs in inclusive early childhood settings requires adaptive teacher strategies rather than uniform instructional responses. These findings reinforce the relevance of creating structured mechanisms through which developmental diversity can be recognized before educational responses are determined.

The same argument extends to the relationship between schools and families. Screening information derived solely from classrooms may overlook behaviors and developmental patterns that are more visible within family environments, while parent observations may not capture demands specific to formal learning contexts. Research involving Indonesian parents of children with special needs demonstrates the importance of recognizing the family as part of the broader developmental ecology surrounding the child (Musri et al., 2024). The present framework's inclusion of both teacher and parent perspectives is therefore potentially significant because it treats developmental identification as information emerging across contexts rather than from a single observer. This does not make different informants interchangeable; instead, discrepancies between their observations may themselves become relevant information for further professional assessment.

The broader relevance of the study also lies in its relationship with inclusive education. Inclusive systems require mechanisms for recognizing learner diversity before developmental difficulties are reduced simply to poor achievement, problematic behaviour, or limited participation. A multidomain framework may help shift attention from isolated classroom performance toward interactions between developmental functioning and educational participation. Recent work on inclusive schools in Indonesia shows that participation by students with disabilities is influenced not only by individual characteristics but also by school policies, teacher roles, social relationships, and institutional support (Supriyanto et al., 2025). This complements Darling-Hammond et al.'s (2020) argument that educational practice should respond to the integrated nature of learning and development. The instrument developed here may eventually contribute to such practice, although its educational effectiveness has not yet been empirically demonstrated.

Methodologically, the study also offers a potentially useful approach for developing screening instruments in culturally diverse contexts. Rather than directly transferring an existing instrument, development began with local users, connected their perceived needs

with the literature, translated constructs into observable indicators, and subjected those indicators to multidisciplinary expert review. This sequence provides a potentially replicable development strategy, although not a universally transferable item set. Boateng et al., (2018) emphasize explicit progression from construct definition and item generation toward content evaluation and psychometric testing. Kyriazos and Stalikas (2018) similarly position standardization as a staged process. Comparable development research in Indonesian inclusive education also demonstrates the value of designing educational models around the characteristics and support requirements of learners with disabilities rather than assuming uniform applicability (Abdillah et al., 2025).

The ICC findings must be interpreted within the same developmental logic. The average-measure ICC of 0.949 demonstrates excellent agreement when seven expert ratings are aggregated, whereas the single-measure ICC of 0.700 indicates good agreement at the individual-rater level. According to Koo and Li (2016), ICC interpretation depends on the statistical model, unit of measurement, and intended use of ratings. The present coefficients therefore strengthen confidence in the consistency of expert evaluation, not in the reliability of scores generated when teachers or parents eventually use the instrument. A content-validated instrument with consistent expert judgments may still perform differently when applied to heterogeneous children in real educational environments. The central psychometric issue therefore shifts from whether experts agree to whether resulting scores behave reliably and meaningfully in the intended population.

Taken together, these findings suggest that early developmental screening in educational settings may benefit from an approach that combines multiple developmental domains with contextual information and clearly distinguishes screening from diagnosis. Such an orientation is consistent with international evidence emphasizing early identification (Kaiser et al., 2024), while Lipkin and Macias (2020) highlight systematic developmental surveillance and appropriate follow-up. Population-specific validation further demonstrates that screening cannot be separated from the cultural and educational contexts in which it is applied (Shum et al., 2022). Recent Indonesian studies similarly connect meaningful inclusion with early recognition of learning challenges and educational responses adapted to learner diversity (Yunitasari et al., 2025; Nurdin et al., 2024). The Indonesian experience therefore has relevance beyond its immediate setting, opening comparative possibilities for culturally

responsive developmental screening within inclusive education while maintaining psychometric rigor and appropriate professional boundaries.

Despite these contributions, several limitations of the study should be acknowledged. The study remains at an early stage of instrument development and has not yet included field-based psychometric validation with the intended population. Therefore, construct validity, internal consistency, test-retest reliability, sensitivity, specificity, and screening accuracy remain unestablished. The purposive sample also limits the representativeness of the findings, while the seven-expert panel cannot fully capture the disciplinary and cultural diversity relevant to developmental screening. These limitations require the findings to be interpreted as preliminary evidence of content validity and consistency among expert judgments rather than as evidence of a fully validated psychometric or diagnostic instrument.

D. Conclusion

This study sets out to develop and validate a multidomain developmental screening instrument for the early identification of neurodevelopmental difficulties in educational settings. The findings provide preliminary support for an instrument that integrates contextual information with cognitive development, behavioral regulation, sensory processing, communication, and adaptive functioning. The needs assessment demonstrated the practical demand for structured and comprehensive screening, while expert validation contributed to substantive refinement of item clarity, specificity, and observability. The resulting content-validation evidence and consistency among expert judgments support the instrument at this initial stage of development, without implying that its broader psychometric properties or diagnostic accuracy have been established.

The principal contribution of this study lies in connecting developmental screening with the educational and ecological contexts in which children's functioning is observed. Rather than positioning teachers or parents as diagnosticians, the proposed framework structures their observations as complementary sources of information that may support early recognition and appropriate referral. Conceptually, this approach brings developmental psychology, child health, educational assessment, and inclusive education into a context-sensitive screening framework. Its broader relevance lies in demonstrating how locally grounded developmental information can be systematically organized without separating developmental characteristics from the family, health, and educational environments in which they occur.

Future research should directly address the limitations identified in this study through field-based validation involving the intended users and child population across broader geographical, educational, age, and sociodemographic contexts. Subsequent studies should examine construct validity, internal consistency, test-retest and inter-user reliability, criterion-related and predictive validity, as well as sensitivity, specificity, and screening accuracy. Larger and more diverse multidisciplinary expert panels would further strengthen content representation. Cross-cultural studies should also examine cultural adaptation and measurement equivalence before the framework is applied internationally. These steps are essential for determining whether the preliminary structure established here can develop into a psychometrically robust and practically reliable screening instrument.

Ultimately, the value of developmental screening lies not in assigning children an early label, but in recognizing developmental concerns early enough to connect them with appropriate educational and professional support.

Acknowledgments

The authors would like to express their sincere gratitude to the teachers and parents who generously participated in the needs assessment and provided valuable insights throughout the instrument development process. We also thank the participating schools for facilitating data collection and supporting the implementation of this study.

Our appreciation is extended to the panel of experts in psychology, special education, early childhood education, and inclusive education for their valuable comments and recommendations during the content validation process. Finally, we gratefully acknowledge the Faculty of Education, Universitas Muhammadiyah Jakarta, for its institutional support and encouragement throughout this research.

Declaration of Competing Interest

The author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Generative AI

During the preparation of this manuscript, the authors used [Copilot] to improve the clarity and readability of the text. Output generated by the tool were carefully reviewed and edited by the authors, who take full responsibility for the content of this

article. All substantive intellectual contributions, including conceptualization, analysis, interpretation of data, and final decisions regarding content, arguments, and conclusions, were carried out solely by the authors. The authors take full responsibility for the integrity, originality, and academic quality of this article.

References

- Abdillah, H. T., Firdaus, E., Syafe'i, M., Budiyanti, N., & Tantowi, Y. A. (2025). Religious Character-Based Inclusive Education in General Course for Difabel Students at University. *Jurnal Pendidikan Islam*, 11(1), 189–203. <https://doi.org/10.15575/jpi.v11i1.44011>
- American Educational Research Association. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.
- Apriliani, I., Pahrudin, A., Koderi, K., & Syafril, S. (2024). Management of Inclusive Education: An Implementation. *Munaddhomah*, 5(1), 112–125. <https://doi.org/10.31538/munaddhomah.v5i1.935>
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. In *Frontiers in Public Health* (Vol. 6). Frontiers Media S.A. <https://doi.org/10.3389/fpubh.2018.00149>
- Boulton, K. A., Hodge, M. A., Jewell, A., Ong, N., Silove, N., & Guastella, A. J. (2023). Diagnostic Delay in Children with Neurodevelopmental Conditions Attending a Publicly Funded Developmental Assessment Service: Findings from the Sydney Child Neurodevelopment Research Registry. *BMJ Open*, 13(2). <https://doi.org/10.1136/bmjopen-2022-069500>
- Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., Vaivada, T., Perez-Escamilla, R., Rao, N., Ip, P., Fernald, L. C. H., MacMillan, H., Hanson, M., Wachs, T. D., Yao, H., Yoshikawa, H., Cerezo, A., Leckman, J. F., & Bhutta, Z. A. (2017). Nurturing Care: Promoting Early Childhood Development. In *The Lancet* (Vol. 389, Number 10064, pp. 91–102). Lancet Publishing Group. [https://doi.org/10.1016/S0140-6736\(16\)31390-3](https://doi.org/10.1016/S0140-6736(16)31390-3)
- Choo, Y. Y., Yeleswarapu, S. P., How, C. H., & Agarwal, P. (2019). Developmental Assessment: Practice Tips for Primary Care Physicians. *Singapore Medical Journal*, 60(2), 57–62. <https://doi.org/10.11622/smedj.2019016>
- Colizzi, M., Lasalvia, A., & Ruggeri, M. (2020). Prevention and Early Intervention in Youth Mental Health: Is it Time for a Multidisciplinary and Trans-Diagnostic Model for Care? In *International Journal of Mental Health Systems* (Vol. 14, Number 1). BioMed Central Ltd. <https://doi.org/10.1186/s13033-020-00356-9>

- Cook, D. A., & Beckman, T. J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. In *American Journal of Medicine* (Vol. 119, Number 2, pp. 166.e7-166.e16). Elsevier Inc. <https://doi.org/10.1016/j.amjmed.2005.10.036>
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for Educational Practice of the Science of Learning and Development. *Applied Developmental Science*, 24(2), 97-140. <https://doi.org/10.1080/10888691.2018.1537791>
- DeVellis, R. F. (2020). *Scale Development: Theory and Applications*. (L. Bickman & D. J. Rog, Eds.; 4th ed.). Sage Publications.
- Ediyanto, E., Kawai, N., Hayashida, M., Matsumiya, N., Siddik, M. A. B., & Almutairi, A. T. (2021). Indonesian Teachers' Attitudes Toward Inclusive Education. *Discourse and Communication for Sustainable Education*, 12(2), 31-44. <https://doi.org/10.2478/dcse-2021-0014>
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational Research An-Introduction: 8th edition* (7th ed.). Boston: Pearson/ Allyn & Bacon.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning. www.cengage.com/highered
- Kaiser, A. P., Chow, J. C., & Baumingham, J. E. (2024). Untying the Gordian Knot of Early Language Screening and Improved Developmental Outcomes. In *JAMA Network Open* (Vol. 7, Number 1). American Medical Association. <https://doi.org/10.1001/jamanetworkopen.2023.54529>
- Khandare, S., & Goyal, A. (2024). Early Detection of Global Developmental Delay in Infants using Indian Screening Tools: Insights from a Tertiary Care Hospital and Implications for Raising Awareness about Neurodevelopmental Outcomes - A Cross-Sectional Study. *International Journal of Health Sciences and Research*, 14(5), 404-413. <https://doi.org/10.52403/ijhsr.20240553>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kyriazos, T. A., & Stalikas, A. (2018). Applied Psychometrics: The Steps of Scale Development and Standardization Process. *Psychology*, 9(11), 2531-2560. <https://doi.org/10.4236/psych.2018.911145>
- Lipkin, P. H., & Macias, M. M. (2020). Promoting Optimal Development: Identifying Infants and Young Children with Developmental Disorders through Developmental Surveillance and Screening. *Pediatrics*, 146(4), 1-20. <https://doi.org/10.1542/peds.2019-3318>

- Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones, E. J. H., Jones, R. M., Pickles, A., State, M. W., Taylor, J. L., & Veenstra-VanderWeele, J. (2020). Autism Spectrum Disorder. *Nature Reviews Disease Primers*, 6(1). <https://doi.org/10.1038/s41572-019-0138-4>
- Mohiyeddini, C. (2024). Cross-Cultural Online Education: Benefits, Challenges, and Solutions. *Online Learning in Educational Research*, 4(2). <https://doi.org/10.58524/oler.v4i2.456>
- Musri, A., Izzaty, R. E., & Gamil, A. (2024). The Effect of Optimism and Gratitude on Subjective Well-Being of Parents of Children with Special Needs. *International Journal of Islamic Educational Psychology*, 5(1), 122–141. <https://doi.org/10.18196/ijiep.v5i1.21548>
- Noordink, T., Tijmstra, J., Verharen, L., Schalk, R., & Van Regenmortel, T. (2025). The Development and Validation of the Psychological Empowerment Scale for Older People. *Journal of Gerontological Social Work*, 68(7), 889–910. <https://doi.org/10.1080/01634372.2024.2441266>
- Nurdin, A., Hendra, Khozin, Haris, A., Zainab, N., & Yahaya, M. Z. (2024). Developing the Islamic Religious Education Curriculum in Inclusive Schools or Madrasah and its Implementation: A systematic literature review. *Jurnal Pendidikan Agama Islam*, 21(1), 94–110. <https://doi.org/10.14421/jpai.v21i1.6907>
- Ratnawulan, T., Ahmad, A. C., Effendi, Z. R., Achmad Syam, R. Z., & Achmad, W. (2025). Teachers' Strategies in Developing Scientific Literacy for Children with Special Needs in Inclusive Early Childhood Education. *Nazhruna: Jurnal Pendidikan Islam*, 8(3), 682–698. <https://doi.org/10.31538/nzh.v8i3.433>
- Roebianto, A., Savitri, S. I., Aulia, I., Suciyan, A., & Mubarakah, L. (2023). Content Validity: Definition and Definition and Procedure of Content Validation in Psychological Reserach. *TPM - Testing, Psychometrics, Methodology in Applied Psychology*, 30(1), 5–18. <https://doi.org/10.4473/TPM30.1.1>
- Romero Jeldres, M., Díaz Costa, E., & Faouzi Nadim, T. (2023). A Review of Lawshe's Method for Calculating Content Validity in the Social Sciences. *Frontiers in Education*, 8. <https://doi.org/10.3389/feduc.2023.1271335>
- Schonhaut, L., Maturana, A., Cepeda, O., & Serón, P. (2021). Predictive Validity of Developmental Screening Questionnaires for Identifying Children with Later Cognitive or Educational Difficulties: A Systematic Review. *Frontiers in Pediatrics*, 9. <https://doi.org/10.3389/fped.2021.698549>
- Shum, K. K. man, Zheng, Q., Wong, W. lap L., Wong, Y. M., & Lam, C. W. C. (2022). Validation of a Developmental Screening Checklist for Chinese Preschoolers in Hong Kong. *Child Neuropsychology*, 28(8), 997–1030. <https://doi.org/10.1080/09297049.2022.2038117>

- Slavec, A., & Drnovšek, M. (2012). A Perspective on Scale Development in Entrepreneurship research. *Economic and Business Review*, 14(1), 3. <https://doi.org/10.15458/2335-4216.1203>
- Supriyanto, S., Imron, A., Maisyaroh, M., Muslihathi, M., Syarif, I., Puspitasari, F. F., & Sulaiman, A. (2025). Social Practices of Students with Disabilities in Inclusive School: A Bourdieusian Analysis of Habitus, Capital, and Field. *Munaddhomah*, 6(4), 635–649. <https://doi.org/10.31538/munaddhomah.v6i4.2055>
- Velikonja, T., Edbrooke-Childs, J., Calderon, A., Slead, M., Brown, A., & Deighton, J. (2017). The Psychometric Properties of the Ages & Stages Questionnaires for ages 2-2.5: A Systematic Review. In *Child: Care, Health and Development* (Vol. 43, Number 1, pp. 1–17). Blackwell Publishing Ltd. <https://doi.org/10.1111/cch.12397>
- Warren, R., Kenny, M., Bennett, T., Fitzpatrick-Lewis, D., Ali, M. U., Sherifali, D., & Raina, P. (2016). Screening for Developmental Delay among Children Aged 1-4 Years: A Systematic Review. *CMAJ Open*, 4(1), E20–E27. <https://doi.org/10.9778/cmajo.20140121>
- Yunitasari, D., Lasmawan, I. W., Arnyana, I. bagus P., & Ardana, I. M. (2025). Breaking Barriers: Advancing Inclusive Education Through Early Identification of Learning Challenges. *Nazhruna: Jurnal Pendidikan Islam*, 8(1), 139–153. <https://doi.org/10.31538/nzh.v8i1.133>
- Zwaigenbaum, L., Bauman, M. L., Stone, W. L., Yirmiya, N., Estes, A., Hansen, R. L., McPartland, J. C., Natowicz, M. R., Choueiri, R., Fein, D., Kasari, C., Pierce, K., Buie, T., Carter, A., Davis, P. A., Granpeesheh, D., Mailloux, Z., Newschaffer, C., Robins, D., ... Wetherby, A. (2015). Early Identification of Autism Spectrum Disorder: Recommendations for Practice and Research: S10-S40. In *Pediatrics* (Vol. 136). <https://doi.org/10.1542/peds.2014-3667C>