Al-Qantara
ISSN: 1988-2955

Received: 08 April 2025
Accepted: 25 June 2025
Published: 18 July 2025

Volume 25, Issue 2, 2025

ASSESSING PRE-LINGUISTIC SKILLS: SCOPING REVIEW OF SCREENING TOOLS AND METHODS

Sana Mukhtar[*1], Dr. Syeda Razia Bukhari[2], Dr. Humaira Shamim[3], Dr. Huma Maqsood[4], Dr. Imran Amjad[5], Dr. Muhammad Imran Tanveer[6], Dr. Asghar Khan[7]

*[*1]PhD Scholar Riphah International University Islamabad/Senior Lecturer and PM BSSS & BS Psychology, Shaheed Zulfikar Ali Bhutto Institute of Science & Technology University Islamabad*
[2]Assistant Professor and Student Counselor, Shaheed Zulfikar Ali Bhutto Institute of Science & Technology University Islamabad
[3,5,7]Professor Riphah International University Islamabad
[4]Assistant Professor Shaheed Zulfikar Ali Bhutto Institute of Science & Technology University Islamabad
[5]Assistant Professor, Curtin University, Sarawak, Malaysia

[*1]sanamukhtar1@gmail.com

DOI: https:/doi.org/ 10.5281/zenodo.16086476

ABSTRACT

Background

Given the importance of early language acquisition and its connection with psychosocial development, various tools have been developed to identify red flags in the first two years of children's lives. However, these tools vary in terms of their use and measurement properties. Therefore, this review examines the measurement properties of scales for assessing pre-linguistic skills to identify early predictors and diagnostic properties of tools and the age range for assessing pre-linguistic skills of 0 to 2-year-old children.

Methodology

Electronic databases including Google Scholar, PubMed, PsycINFO, ASHA, CINAHL, ERIC, Cochrane Library were searched, along with citation-chaining strategies. An initial pool of 50 tools was identified. Based on inclusion and exclusion criteria, 15 tools were retained for further review. A total of 558 articles were identified, from which 52 were reviewed for psychometric properties of the selected tool.

Results

The findings reveal several limitations across assessment tools. These gaps underscore the need for a more comprehensive and inclusive tool for assessing pre-linguistic skills.

Conclusion

Based on the review, we recommend the development of a web-based tool for assessing pre-linguistic skills in infants. Such a tool would be enhancing early identification and intervention, which is crucial for effective intervention and optimal developmental outcomes.

Keywords: *systematic review[1], pre-linguistic skills[2], assessment tool[3], language acquisition [4], psychometric properties[5]*

Al-Qantara
ISSN: 1988-2955

INTRODUCTION

Evidence on early predictors of neuro-developmental delays emphasizes the importance of careful monitoring of early social communication skills, including emotional expression, eye contact, attention, command following, pointing, turn-taking, and play skills (Brewer et al., 2020; Parmeggiani et al., 2019). Assessment of these skills requires a screener or assessment tool that measures these constructs (Delehanty et al., 2018). During the pre-linguistic communication period, which begins at birth and extends until a child's first meaningful utterances, children communicate with a variety of functions, starting with eye gaze and social-emotional expressions, and later using gestures and other nonverbal communication styles (McLeod et al., 2017). These skills form the foundation for later use of vocabulary, syntax, and communication (Keen et al., 2016).

Language acquisition is a precursor for the development of social, emotional, and behavioral components required for a balanced personality, and delay in language acquisition leads to problems in socio-emotional adjustments in life (Laws & Hall, 2014). Therefore, parents and clinicians particularly focus on children's vocal skills during the preschool years and beyond, especially when dealing with neuro-developmental disorders such as autism spectrum disorder (ASD). Research indicates that early diagnosis, followed by effective intervention based on pre-linguistic skills such as eye gaze, attention, gestures, and vocalization, can enhance the language skills of children (DiStefano & Kasari, 2016).

Therapists rely on both standardized and informal assessment methods for children's language and cognitive skills assessment (Ward et al., 2021). Standardized assessments are a useful tool for diagnosis and early intervention planning, as well as predicting future performance in various domains, including language and behavior (Flax et al., 2009; Tambyraja et al., 2017). However, standardized tools pose problems such as floor effects, validity issues, and difficulties in performance related to neuro-developmental disorders (Kaiser & Roberts, 2011; Paul, 2007), which emphasize the need to develop a holistic tool that focuses on the early language skills of infants.

Clinicians often rely on parents' and caregivers' reports on children's skills, but these assessment reports have validity issues due to the varied assessment criteria of caregivers (McCoy et al., 2018). Hence, UNICEF has defined child development standards for different countries (Clark et al., 2020).Country-specific standards provide a common platform for stakeholders to initiate appropriate action plans and strategies for child development (Fernald et al., 2017).

Based on the background and problem, our study presents a review of the pre-linguistic skills assessment tools. It offers an overview of the psychometric properties of the assessment tools based on the domains of pre-linguistic skills to identify salient features of the 15 assessment tools reviewed in this study. Thus, it provides a baseline for the development of a pre-linguistic skills assessment tool that is holistic and comprehensive and covers all domains of pre-linguistic skills.

This review is related to reliability, validity, development norms, translated versions, domains covered, specificity, and sensitivity of different tools. Therefore, it is relevant to researchers and practitioners keen on the assessment of language acquisition, pre-linguistic skills, or early language development phases.

We proceed with the following objectives:

1. To find out the tools available for assessment of pre-linguistic skills of infants
2. To appraise, compare, and summarize the quality of measurement properties of all pre-linguistic skill-assessment tools
3. To identify areas of improvement in pre-linguistic assessment of 0 to 2-year-old infants.

Methodology

Protocol

This review covers all relevant tools used for pre-linguistic assessment until April 2023, follows the guidelines from the Cochrane diagnostic test accuracy protocol (Deeks et al., 2013) and the COSMIN guidelines for tool reviews (Mokkink et al., 2018). The methodology was also specified in a protocol published on the PROSPERO database for systematic

reviews (CRD42021268445) and was further improved through discussions with subject matter experts.

Search Strategy

We had extended discussions with clinical psychologists and speech-language pathologists and conducted an extensive literature review for the selection of scales and pre-linguistic assessment tools. We specifically considered the scales that are currently being used by Clinical Psychologists and Speech-Language Pathologists. In literature, various databases were utilized to find the reviews of Early childhood development screeners, e.g., web of science, Scopus, google scholar etc. Moreover, we used various search engines to find reviews of early childhood development screeners. The search was done using the terms: pre-verbal skills, pre-verbal skills assessment, tools used by speech and language pathologists, pre-linguistic skills, early language development, and language acquisition in early years. Moreover, we organized the search in the following ways:

Step 1: Search related to pre-linguistic skills combined with "OR"

Step 2: Search all keywords related to early language development "OR"

Step 3: Search all keywords related to pre-speech skills with "OR"

Step 4: Search all keywords with early language acquisition

Step 5: Combine all the above searches with "AND"

Data Sources

We systematically searched different electronic databases such as Google Scholar, PubMed, PsyhcINFO, ASHA, CINAHL, ERIC, Cochrane Library databases, 5 books on child development (Berk, 2020; Levey, 2019; Santrock, 2014), and 1 manual of the scale. To record the screening process for transparency in all stages, we used Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines. (Figure 1 & 2)

Data Extraction

Initially, 50 assessment tools were identified, and 23 met the eligibility criteria. Relevant abstracts and psychometric properties of the scale were explored to additionally filter the tools. The inclusion of tools had the following criteria:

1. Tools measuring pre-linguistic skills
2. Tools with reported measurement properties
3. Tools covering the age range from birth onwards
4. Tools that help plan early intervention plans
5. Tools that are published and can be purchased
6. Tools only published and standardized in the English language
7. Standardized tools being used by speech/language pathologists and psychologists

In the first step, 50 assessment tools were selected from the literature but 15 were retained based on inclusion and exclusion criteria. In the second step, 558 articles were reviewed for the psychometric properties of the selected assessment tools. All the relevant articles on selected assessment tools were reviewed. At the initial stage, 168 were chosen because of the information provided on measurement properties. From the remaining 168 articles, 52 met the criteria as they had relevant information on the selected tools.

Twenty-seven tools were excluded as they did not cover the pre-linguistic skills domain, the age from birth, and were not standardized in English. For the remaining 23 scales that met the eligibility criteria, we created a spreadsheet to record their measurement properties.

Eligibility Criteria for selection of articles:

The inclusion criteria were:

1. Peer-reviewed articles published at the time of test development (till April 2023)
2. Articles with descriptions of test construction, reliability, validity, sensitivity, specificity, standardization sample, and limitations of the tool
3. Articles with psychometric properties of the assessment tools

The exclusion criteria were:

1. Articles not accessible through open source

Data Analysis:

To critically appraise, compare, and summarize the quality of measurement properties, thematic analysis method was applied to achieve the research objectives (Braun & Clarke, 2006). Moreover, Saldana (2018) recommended thematic analysis as compared to content analysis when specific and precise meanings within a given context are required. Similarly, it was also evidenced that thematic analysis tends to build the social construction as well as building the perspectives from the data (Braun & Clarke, 2013). Therefore, the views of practitioners were identified using thematic analysis for perspective building. We employed thematic analysis as prescribed by Braun and Clarke (2006) through the six stages which starts with familiarity with data, followed by generating initial codes. After generating the codes, these were transformed into themes. In the fourth stage, researchers reviewed the themes so that themes were defined and named in stage five. Finally, in the last stage, report generation was initiated on the basis of findings from the themes. However, stage two, three and four were an iterative process for data reduction from codes and themes.

Results

The 15 selected tools were organized in Excel to record information for each assessment tool. Different research articles were read to gather further information on the tools and report the measurement properties of each tool. The initial pool identified 558 articles. At the initial stage, 168 articles were selected because of the relevant information on measurement properties. From the remaining 168 articles, 52 met the criteria of relevant information on the selected tools. Reaching a saturation point when most of the detail on psychometric properties was obtained, further review was stopped.

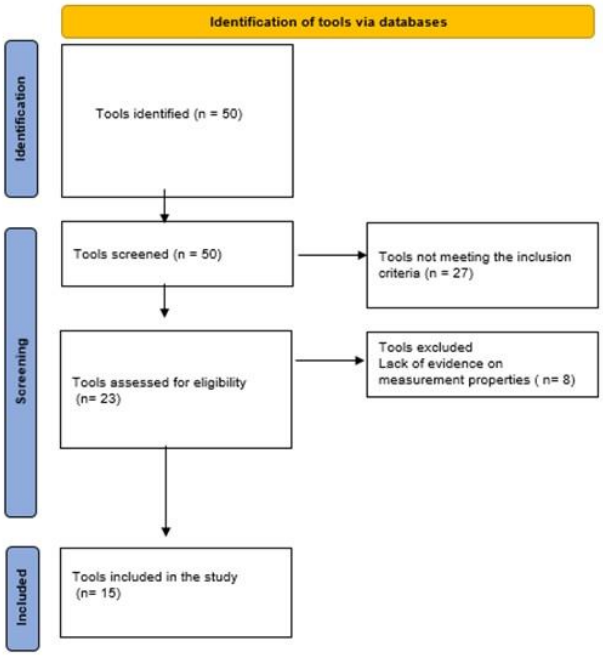

Figure 1 PRISMA flowchart for selection of tools

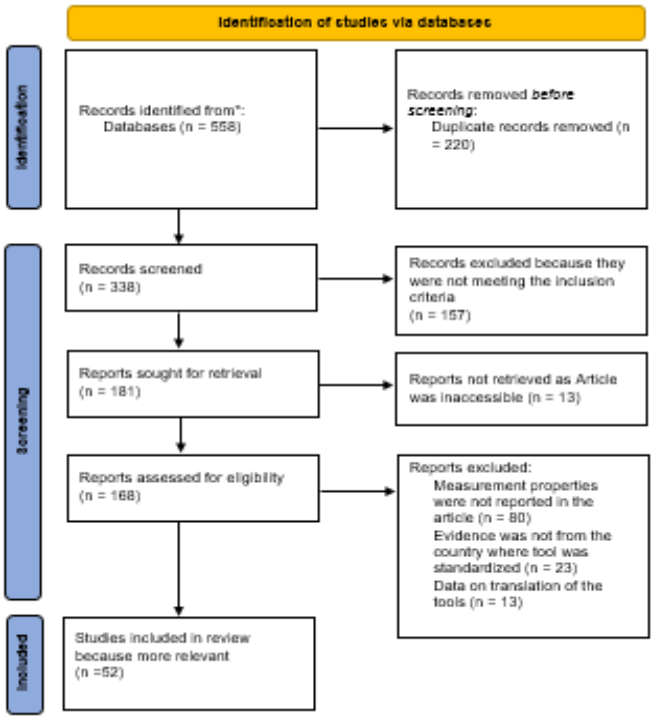

Figure 2: Flow chart for articles on selected tools

Discussion

The selected assessment tools varied distinctly in their psychometric properties and the domain of pre-linguistic skills as discussed in the subsection below.

1.1 Ages and Stages Questionnaire (ASQ):

The Ages and Stages Questionnaire (ASQ) was developed by Squires and Bricker (2009). ASQ is considered the norm-referenced test and is translated into different languages such as Chinese, French, Arabic, Turkish, Persian, Norwegian, Spanish, Vietnamese, Dutch, Hindi, and Persian (Bian *et al.*, 2012; Vameghi *et al.*, 2013; Singh *et al.*, 2017; Charafeddine *et al.*, 2019; Rousseau *et al.*, 2021). It covers five domains: interpersonal skills, motor learning, fine motor skills, problem-solving, and receptive and expressive communication at different ages. The scale includes 21 forms and scoring sheets for different age stages up to 60 months (Squires et al., 2009). It was developed with a standardized sample of 15,138 infants, and it mainly addresses overall development progress (Squires et al., 2009). The ASQ-3 has been reported to have a validity range of 0.82 to 0.88, test-retest reliability of 0.91, and inter-rater reliability of 0.92 (Glascoe *et al.*, 2012). Its internal consistency ranges from 0.51 to 0.87 and sensitivity and specificity are reported to be 86% and 85%, respectively (Singh et al., 2017; Squires, J., & Bricker, 2009).

The ASQ also has a Social-Emotional 2nd edition, which was developed in 2015 to cover social-emotional behaviors of self-regulation, social communication, compliance, adaptability, autonomy, affect, and social interaction (Squires et al., 2002).

Limitations

ASQ is screening tool and not a diagnostic tool. It is based on Western developmental norms so it decreases its applicability for children from diverse cultural or socio-economic backgrounds (Bian *et al.*, 2017; Chen *et al.*, 2017). This also increase the risk of false positive as parenting style and child trait can impact the final scores. The sensitivity and specificity also vary from study to study (Letts et al., 2023). The studies have also reported reliability issues for some

subscales (Martinussen & Valla, 2013). The parental reports are often biased and the information provided by parents is inaccurate. Demographics of parents also has huge impact on total score of the ASQ questionnaire.

While the ASQ is a practical tool for early developmental screening, its limitations highlight the need for complementary assessments and careful consideration of cultural and contextual factors. These limitations suggest that while the ASQ can be effective, it should not be the sole method of assessment, and results should be interpreted with caution.

1.2 Bayley Scales of Infant and Toddler Development (III & IV) – 3rd & 4th Edition

The Bayley Scales of Infant and Toddler Development is a parent-based assessment tool co-authored by Nancy Bayley and Glen Aylward in its 4th edition (Del Rosario et al., 2021). It assesses various aspects of development including cognitive, linguistic, motor, social-emotional, and adaptive behavioral skills of 16 days to 42 months old children. It can be administered either through paper-and-pencil or an online testing platform, Q-global (Albers & Grieve, 2007).

Bayley-III assesses children from birth to 42 months of age, while Bayley-IV assesses children from 16 days to 42 months of age (Bayley, 2006; Alfonso et al., 2022.; Milne et al., 2012). Bayley-III takes between 45-90 minutes and provides detailed information about a child's development (Bayley, 2006); whereas, the administration time of Bayley-IV is between 15 to 25 minutes (Bayley & Aylward, 2019).

The Bayley-III assessment tool includes a Cognitive scale with 91 items, a receptive language scale with 49 items, an expressive language scale with 48 items, a fine motor scale with 66 items, and a gross motor scale with 72 items. The social-emotional scale is derived from the Greenspan chart, and the adaptive behavior scale is derived from the Adaptive Behavior Assessment System (ABAS) (Bayley, 2006; Anderson & Burnett, 2017).

Bayley-IV, the updated version of the tool, has 81 items for the cognitive scale, 42 items for receptive language, 37 items for expressive language, 46 items for the fine motor scale, and 58 items for the gross motor scale. The social-emotional scale is the same as in Bayley-III while the adaptive behavior scale uses the Vineland Behavior Assessment System (Bayley & Aylward, 2019).

Norms of Bayley-III were developed with 1700 typically developed children (Anderson & Burnett, 2017b). For Bayley-IV, normative data was collected between 2014 and 2016, using a sample of 1,753 children from diverse regions, ethnicities, and socioeconomic backgrounds in the United States. The sample included term and preterm infants and toddlers, as well as children with various medical comorbidities. Bayley-IV norms are, therefore, based on a more recent and diverse sample compared to Bayley-III. These norms provide a reference point to interpret Bayley scores and evaluate a child's developmental progress (Bayley, 2019).

The internal consistency for the domains of cognitive, language, and motor skills are 0.93 to 0.95, 0.85 to 0.91 for social-emotional, and 0.91 to 0.98 for adaptive behavior. The technical manual of Bayley reports internal consistency coefficients (Cronbach's alpha) that range from 0.65 to 0.95 for cognitive, language, and motor scales, and from 0.56 to 0.89 for social-emotional and adaptive behavior scales. The test consistency for cognitive-motor and language is 0.81 to 0.84; whereas, for adaptive behavior, it is 0.72 to 0.87. The interrater reliability is 0.67 to 0.81. The classification accuracy for delays in development is 0.82 and language delays 0.89 in the BSID 4th edition (Piñon, 2010).

Bayley-IV is based on the theory of the neuro-environmental model of development (Floyd et al., 2015; Alfonso et al., 2022) and includes new normative data based on a more diverse and representative sample of children (Bayley & Aylward, 2019; Alfonso et al., 2022).

Limitations

The Bayley III edition has faced criticism for overestimating developmental abilities of children, thus considered ignoring children who need intervention. This is more evident in cognitive, language and motor domains (Anderson & Burnett, 2017a). Studies also indicate that Bayley ha limited

predictive validity for intelligence or academic performance. The test retest reliability indicated inconsistencies thus questioning the reliability of the scale especially with children in early years (Koshy et al., 2024). Test is also time consuming, expensive and require professional training to administer the scale. Moreover, the scale is standardized for specific population and not suitable for children from diverse backgrounds.

1.3 The Battelle Developmental Inventory (BDI) 2nd and 3rd edition

The Battelle Developmental Inventory (BDI) was initially developed by Jen Newborg and colleagues at the Battelle Memorial Institute in 1983 (Cunha et al., 2018b). It was designed to evaluate the developmental skills of children from birth to 8 years of age, with emphasis on the early identification of developmental delays and disabilities (Newborg & Company, 2005).

The BDI is one of the assessments used to monitor the five developmental domains outlined in the Individuals with Disabilities Education Act (IDEA) Part C (Berls & McEwen, 1999; Macy *et al.*, 2015; Stone-MacDonald *et al.*, 2018). These domains are further divided into subcategories: Adaptive, Personal-Social, Communication, Motor, and Cognitive, which can be administered in both English and Spanish languages (Encyclopedia of Child Behavior and Development, 2011). The 3rd edition of the BDI, developed in 2019, includes a new domain on social-emotional development to evaluate emotional regulation, social relationships, and play skills (Kilburn, 2020).

The normative sample for the BDI-2 consisted of 5,000 children from the United States, and the BDI-3 sample consisted of over 6,000 children from the United States and Canada. Both editions were normed on different demographic variables such as gender, race/ethnicity, parental education, and geographic region (Newborg & Company, 2005). Their reliability ranges from 0.98 to 0.99; test consistency is above 0.80; and validity is reported to be 0.63 for receptive BDI-2 and 0.73 for expressive. The total domain score on validity is 0.72, and cut-off scores indicate acceptable sensitivity and specificity (Bliss, 2007; Stone-MacDonald, Pizzo and Feldman, 2018). However, scarce evidence is available on the reliability and validity of BDI-3, yet according to the publisher's website, the BDI-3 has internal consistency reliability coefficients ranging from .73 to .98 for different subtests and test-retest reliability coefficients ranging from .60 to .99 (Newborg, 2020).

Limitations

As noted by several researchers one potential drawback of the BDI is its lengthy administration time (90 minutes), as well as the time required to score the assessment. To address this issue, the Battelle Development Screening Test was developed as a quicker assessment tool for identifying developmental delays in children. While the BDI has primarily been used in American populations, several studies have highlighted its potential for effective use in different cultural contexts (Cunha et al., 2018a). However, it is noteworthy that neither the BDI 2nd nor 3rd edition covers emotional regulation or executive functioning.

1.4 The Vineland Adaptive Behavior Scale 3rd Edition

The Vineland Adaptive Behavior Scale 3rd edition, developed by Sara et al. (2016) is an updated version of the Vineland Adaptive Behavior Scale 2nd edition (Sparrow & Saulnier, 2016). This scale measures adaptive behaviors from birth to 90 years of age and assesses skills such as communication, daily activities, and interactional behaviors. The subdomains of the scale consist of receptive, expressive, written, personal, domestic/numeric, school-related activities, social interaction, play skills, problem-solving abilities, and gross and fine motor skills.

The normative sample for the parent/caregiver form was 2,560 individuals, and for the teacher form, it was 1,415 individuals. The total administration time for the scale ranges from 10 minutes for parent/caregiver forms to 40 minutes for comprehensive interview forms and can easily be administered by a therapist with a Master's degree (Sparrowtti & Saulnier, 2016; Pepperdine & McCrimmon, 2018).

Internal consistency of the Vineland Adaptive Behavior Scale 3 ranges from 0.94 to 0.99, and the reliability for all domain-level forms and adaptive behavior composite is between 0.86 to 0.97

(Pepperdine & McCrimmon, 2018). The test-retest reliability for the comprehensive form adaptive domains and adaptive behavior composite is from 0.64 to 0.94 across all ages and 0.62 to 0.92 for all domain-level adaptive domains and adaptive behavior composites. The interrater and inter-interviewer reliability for comprehensive form adaptive and adaptive behavior composite is between 0.61 and 0.87, and the range is between 0.58 and 0.93 for all domain-level adaptive domains and adaptive behavioral composite. The concurrent validity range is from 0.32 to 0.83 (Sparrowtti & Saulnier, 2016; Farmer et al., 2019).

Limitations

Despite various strengths of scale, there are inconsistencies between subscales and interrater reliability variability (Farmer et al., 2019). Moreover, the Vineland Adaptive Behavior Scale 3 is representative of the US population, so there is a possibility of cultural biases that could affect the objectivity of the results. In addition, it relies on parents' or caregiver's information. Hence, there are chances of bias and inaccuracies (Kanne et al., 2011). Therefore, the scale is not suitable for children with ineffective communication skills (Carter *et al.*, 2004; Volkmar *et al.*, 2014). Most researchers or therapists prefer online scoring due to the time-consuming manual scoring. However, online scoring has additional financial costs (Carter *et al.*, 2017).

1.5 Brigance Early Childhood Screens III

According to research by Glascoe et al. (1992), the Brigance Early Childhood Screen-III is an effective tool for diagnosing developmental delays in children aged from birth to seven years of age. The tool includes different screens for infants, toddlers, early preschoolers, and preschoolers. It consists of four assessments: the Early Child Development Inventory (ECDI); Infant and Toddler Developmental Inventory (ITDI); Preschool Developmental Inventory (PDI); and the Comprehensive Inventory of Basic Skills (CIBS II) (*Brigance® Early Childhood Screens III*, n.d.; Dockrell *et al.*, 2017).

The Brigance Early Childhood Screen-III was first developed in the 1980s. Later, after extensive research of over 25 years and user requests, the Brigance Infant and Toddlers' Screener was developed to assess development delays in the first few years of children's lives. The screens are criterion-referenced and can be scored in 15 minutes. Moreover, they diagnose children with language delays, difficulties, learning problems, and delays in cognition (Glascoe, 2002; French, 2013; Pyle *et al.*, 2018).

The screens have been standardized on a sample of 11,000 children selected through stratified sampling based on ethnic background, topographies, and socioeconomic status (Martin & Whiteman, 2017). The sensitivity of the screens ranges from 82 to 86%, and the specificity is 84%, indicating high accuracy. The internal consistency and correlation for test consistency reliability range from 0.84 to 0.99, and inter-rater reliability ranges from 0.90 to 0.99 (Glascoe, 1997).

The validity of the screens is also high, with content validity based on subject matter experts and extensive literature on early childhood development (Martin & Whiteman, 2017). Factor analysis indicates strong construct validity for the items, and concurrent validity is highly correlated to measures of motor development, language, and academic skills, ranging from 0.66 to 0.97. Furthermore, the screens have a high degree of predictive and discriminate validity (Mantzicopoulos, 1999; Pyle et al., 2018).

Limitations

Various researchers recommend cross-validation, longitudinal studies, and the use of screens on children with diverse cultural backgrounds (Brigance, 2016). However, to obtain a complete picture of development, it is necessary to rely on multiple assessments (Janus & Offord, 2007). Moreover, it is noteworthy that the screens are available in only English language and may not be suitable for bilingual children or those who speak other languages (French, 2013).

1.6 Denver Development Screener

The Denver Developmental Screener is widely used for assessing child development from birth to 6 years of age. It was first developed in 1967 and has been

standardized over the past 23 years (Frankenburg and Dodds, 1967; Sperhac and Salzer, 1976). The screener was standardized with a population of 2096 children in Denver; the test-retest reliability is reported to be 95.8%; and the inter-rater reliability is 90% (Frankenburg *et al.*, 1992).

Limitation

However, research conducted by Thomson Delmar Learning (2007) reported that the screener has a high sensitivity rate of 83% yet limited specificity. Another limitation of the Denver Developmental Screener is its cultural bias. Therefore, it is recommended to conduct cross-validation and longitudinal studies to ensure that the screener is appropriate for children with diverse cultural backgrounds (Luiz et al., 2004).

1.7 Development Assessment of Young Children (DAYC) 2nd edition

The Development Assessment of Young Children 2nd Edition, developed by Voress, Maddox, and Hammill, is a norm-referenced test covering five domains: Adaptive behavior, cognition, communication, social-emotional development, and physical development (Voress & Maddox, 2013). It consists of a total of 88 items in the cognitive domain, 78 items in the communication domain, 63 items in the social-emotional domain, 87 items in the physical development domain, and 64 items in the adaptive behavior domain (Voress and Maddox, 2013; Melissa D Swartzmiller, 2014)(Swartzmiller, 2014b).

The test takes 50 to 100 minutes to administer, with each subtest taking 10-20 minutes. The sample for norms was 1832 children from the US, and data was collected for 2009-2011. The reliability for the domains is from 0.80 to 0.98, test-retest reliability for domains is 0.70 to 0.91, and for subdomains, it is 0.82 to 0.90 (Swartzmiller, 2014a). The criterion predictive validity was evaluated by comparing it with the Battelle Development Inventory 2nd Edition and the Development Observation Checklist System 2nd Edition (Gannon, 2007).

Limitations

The accuracy, sensitivity, and specificity for language delay were above 0.70 (Voress & Maddox, 2013). However, the DAYC 2nd edition may offer inadequate results on domains related to infants under 9 months of age, and it has an item gradient that is different from other measures used to predict developmental delays in children (Swartzmiller, 2014b).

1.8 Parents' Evaluation of Developmental Status (PEDS)

PEDS is an evidence-based tool for detecting delays in child development or problems in the behavior of children from birth to 7 years of age. It was developed by Dr. Frances Page Glascoe in 1998 and consists of 10 items to be filled by parents (Glascoe, 2000; Chunsuwan *et al.*, 2016). The domains covered in PEDS are expressive, receptive language, fine and gross motor skills, global/cognitive, self-help, and socio-emotional (Woolfenden et al., 2014). PEDS is translated into 32 different languages such as Spanish, French, Chinese, Portuguese, German, Arabic, Italian, Korean, Thai, Turkish, Dutch, Danish, Norwegian, Swedish, and Finnish (Chunsuwan *et al.*, 2016; Copeland, 2017).

The normative Sample of PEDS includes 1,000 developing children aged 0-8 years from diverse ethnicities and socioeconomic backgrounds (Glascoe, 2000). The PEDS test consistency reliability is 0.98 to 0.99 and inter-rater reliability is 0.82 to 0.92. The assessment is also highly valid and the correlation ranges from 0.90 to 0.99. The sensitivity is 0.79 and the specificity is 0.89, which is considered moderate (Brothers et al., 2008; Vameghi et al., 2015).

Limitations

PEDS is not a diagnostic but a screening tool. Therefore, if any red flags are identified by PEDS, it is necessary to conduct a detailed assessment to confirm the concerns (Glascoe, 2005). Besides, PEDS has subjective nuances that may not consider cultural contexts (Coghlan *et al.*, 2003; Mukherjee *et al.*, 2022; Sheel *et al.*, 2023).

1.9 Mullen Scales of Early Learning (MSEL)

The Mullen Scale of Early Learning (MSEL) was developed by Dr. Eileen Mullen in 1995 as a norm-referenced test that measures visual acuity, gross and

fine motor skills, and receptive and expressive language of 0-5-year-old children (Mullen, 1995; Dumont *et al.*, 2014). MSEL has been translated into Spanish and four South African languages (Bornman *et al.,* 2018; Connery *et al.,* 2019; Colbert *et al.,* 2020). The scales are also adapted in other cultures (Milosavljevic *et al.,* 2019; Nimkar *et al.,* 2021).

The standardization sample includes 1849 children from birth to 68 months of age (Mullen, 1995), and it is considered highly effective in the assessment of children with ASD.

Research on convergent and divergent validity of the MSEL indicates strong standardized scores such as 0.89 for visual reception, 0.84 for fine motor, 0.84 for fine motor, 0.92, and 0.89 for receptive and expressive language (Bishop et al., 2011; Burns et al., 2013).

Limitations

MSEL has been reported to diagnose neuro-developmental conditions but with limited criterion ability across different clinical groups. Therefore, additional research is needed to enhance its sensitivity (Swineford et al., 2015). Moreover, the test is relatively expensive and excludes children with disabilities from standardization. Besides, there is limited evidence of concurrent, content, and construct validity of the test (Bedford et al., 2013).

1.10 Rosette Infant and Toddler Scale

The Rosette Infant and Toddler scale was initially published in 1990 by Louis Rosette, with minor revisions made in later versions in 2005 and 2006. It was meant to fill the gap in the assessment of pre-verbal and verbal skills in children up to 3 years of age. The scale has also been translated into English and Spanish languages (Moon, 2019). It assesses six domains: caregiver interaction and attachment; use of language in social settings; gestures; play skills; understanding of language; and production of language (Ritvo, 2013).

Limitations

The Rosette scale is a criterion-referenced test and was not specifically developed for screening Autism Spectrum Disorder, but it can be used as part of test batteries for assessment of children with ASD (Rossetti, 2006). Moreover, it is an effective tool for assessing developmental delays and determining mild, moderate, and severe levels of communication delays. Nonetheless, the psychometric properties of the Rosette Infant and Toddler scale are not reported in the manual. Therefore, it cannot be used for diagnostic purposes (Hutchins, 2021).

1.11 Kent Inventory of Development Skills 3rd Edition

Kent Inventory of Development Skills was developed by Reuter, Katoff and Wozniak, (2000). It is a parents-based assessment and consists of items covering motor, self-help, cognitive, communication, and social skills. KIDS consists of 252 items and takes 45 minutes to administer (Schludermann, 2001). It can be used effectively with populations of the Netherlands, Spain, Russia, Hungary, and other English-speaking countries. It is paper-based, but PC-based software is also available.

The scale can be used for infants from birth to 15 months of age and is useful for children with severe developmental delays of age 6 years. The scale is standardized with 706 infants. The internal consistency is 0.95, and the ranges for test-retest reliability are 0.86 and 0.98. Scale validity is 0.95, and for 5 domains it is between 0.80 and 0.88.

Limitations

KIDS' parent-based assessment makes it subject to inaccuracies, especially if parents have limited knowledge of child development. Another limitation is that it is standardized with a small sample size which makes its reliability and validity susceptible to application in a generalized population. Furthermore, it covers only infants from birth to 15 months of age and is useful only for children with severe developmental delays up to age 6 years of age (Schludermann, 2001).

1.12 The Ounce Scale

The Ounce scale is an observational tool that is administered through parents. It is a norm-referenced test designed for from birth to 41 months of age and is also translated into Spanish language (Meisels et al.,

2010a). The scale covers six areas of development including interpersonal relationships, self-awareness, peer relationships, language skills, analytical skills, and motor coordination (Ringwalt, 2008b). In addition to monitoring the biological and psychological aspects of growth, the Ounce scale provides guidelines to caregivers as well (Meisels et al., 2010b). The reliability of the Ounce scale ranges from 0.19 to 0.89, and several studies have shown that it is up to 70% accurate in screening children at risk (Ringwalt, 2008a)

Limitations

However, the limited data used in several studies highlights generalizability issues in its use. Another limitation is its subjective nature because of parent-based administration. Furthermore, although it screens for further evaluation, it does not diagnose developmental delays.

1.13 Developmental Profile 3 & 4 (DP-3)

The DP-3 was developed by Alpern (2007), and the latest version, DP-4, was developed in 2020. The Development Profile measures five domains of development: physical development, adaptive skills, social development, emotional development, and cognition and communicative development of children. The latest version, with 190 items, is appropriate for assessing developmental aspects from birth to 21.11 years of age. The scale is translated into Bulgarian, Czech, Danish, and Spanish languages, and administration of this test takes 20-40 minutes, with parent/caregiver interview forms, checklist forms, teacher checklist forms, and clinician rating forms (Ringwalt, 2008a).

DP-4 has an internal consistency of 0.80-0.97, and construct validity is confirmed through inter-scale correlations, factor analysis, and Rasch analyses(*Developmental Profile 4 Manual*, 2020). Its specificity is 0.99 and sensitivity is 0.56. Standardization was done on a sample of 2216 individuals, and the clinical sample included 348 children with a diagnosis of Autism Spectrum Disorder, Attention-Deficit Hyperactivity Disorder, children with intellectual, learning, physical, and other disabilities, hearing-impairment, psychiatric issues, speech and language-related issues, and visual impairment (Alpern, 2020).

Limitations

DP-4 cannot be used for diagnostic purposes alone; rather, it should be used with other formal assessments, parent/caregiver interviews, history of the child, and observation of a child. The sensitivity of DP4 is low, which implies that it may not detect development delays or disorders. Furthermore, it lacks relevant evidence from different ethnic backgrounds. Therefore, administrators need professional competence to complete the test.

1.14 Preschool Language Scale 5

PLS-5 is a professional assessment scale with an English version developed in 2011 and a Spanish version developed in 2012 by Zimmerman, Steiner, and Pond (Restrepo & Silverman, 2001). It is both norm-referenced and criterion-referenced and covers attention, play skills, gestures, vocalization, social communication, semantic understanding, receptive and expressive language as well as emergent literacy skills in children from birth to 7.11 years of age (Zimmerman et al., 1979). The scale is standardized with a sample of 1400 children, and norms are also explained for different age intervals (Wiig et al., 2014). It takes approximately 45-60 minutes to administer (Riley et al., 2019).

PLS-5 is translated into Turkish. However, the previous versions have been translated into several languages (Sahli & Belgin, 2017). The reliability of the scale ranges from 0.86 to 0.95; inter-rater reliability is 0.96 to 0.99; and inter-item consistency is 0.91 and 0.98. The sensitivity of the scale is reported to be 0.83 and the specificity is 0.78 (Riley et al., 2019; Lyons, 2021).

Limitations

Supporting evidence is missing regarding the accurate validity and reliability of PLS-5. Although the scale is widely used, cultural biases, language barriers, and children's interaction with the objects of interest raise several concerns. The scale does serve screening purposes, but for diagnosis, it needs to be used in

combination with other batteries (Lee Zimmerman et al., 2012).

1.15 Receptive-Expressive Emergent Language Test (REEL 3rd edition)

REEL 3rd edition as a norm-reference test was developed by Kenneth et al. (2003). It measures both receptive and expressive language from birth to 3 years of age (Technical, 2021). Its reliability scores are above 0.90. The test-retest score is reported to be consistent, but specificity and sensitivity are not reported. Norms were developed on 1,112 infants and toddlers from 2001 to 2002. The test consists of 60 items, but they are divided into different age groups. It takes 20 minutes to administer the test (Bzoch et al., 2003a).

Limitations

REEL is limited to only language scores and most of the evidence highlights parents' difficulty in understanding the items. The test is in a Yes/No format, which makes it restricted to only what is asked rather than what parents would like to share (Bzoch et al., 2003b).

Conclusion

The findings underscore that all pre-linguistic skills should be covered on a single scale to assess the pre-linguistic skills. Most of the scales were initially developed to assess the stage of language acquisition in early childhood, but the initial versions were improved based on research findings. Further different editions were developed for most of the scales to cover social and emotional domains, which highlight the need for a holistic tool that exclusively covers pre-linguistic skills. Hence, it is concluded that we need to develop a tool for pre-linguistic assessment skills that covers all pre-linguistic skills including reflexes, eye contact, attention, intent, play skills, social referencing, gestures, and emotions. Moreover, this tool must also adopt a standard approach to define the specific domains of child development, such as cognition, language, cultural knowledge, social-emotional development, language acquisition, motor development, and emerging literacy.

Theoretical and Practical Implications:

The current review has utilized "theory of change". The focus is not on "type of scales used for assessment of child language but to get a deeper understanding of "what are the best scale, why they are considered good and what can be done to develop a future assessment?".The review has helped in defining the taxonomy of pre-linguistic skills. In addition, it facilitated in developing a tool that is culturally appropriate and covering the complex nature of pre-linguistic skills. The new scale is based on three domains of Pre-linguistic skills: the foundational sensory motor skills, social interaction and communication skills and Emotional and expressive skills and give more comprehensive evaluation of a child. This new scale is using both Artificial intelligence and Clinician decision making for a better intervention.

REFERENCES

Albers, C. A., & Grieve, A. J. (2007). Test Review: Bayley, N. (2006). Bayley Scales of Infant and Toddler Development– Third Edition. San Antonio, TX: Harcourt Assessment. *Journal of Psychoeducational Assessment*, *25*(2), 180–190. https://doi.org/10.1177/0734282906297199

Alfonso, V. C., Engler, J. R., & Turner, A. D. (n.d.). OVERVIEW OF THE BAYLEY–4. *Essentials of Bayley-4 Assessment*, 1.

Alpern, G. D. (2007). Developmental profile 3 (DP-3). *Los Angeles, CA: Western Psychological Services*.

Alpern, G. D. (2020). *"DP-4: Developmental Profile*. Western Psychological Services Torrance California.

Anderson, P. J., & Burnett, A. (2017a). Assessing developmental delay in early childhood—concerns with the Bayley-III scales. *The Clinical Neuropsychologist*, *31*(2), 371–381.

Anderson, P. J., & Burnett, A. (2017b). Assessing developmental delay in early childhood — concerns with the Bayley-III scales. *Clinical Neuropsychologist*, *31*(2), 371–381. https://doi.org/10.1080/13854046.2016.1216518

Bayley, N., & Aylward, G. P. (2019). *Bayley Scales of Infant and Toddler Development (4th ed.) technical manual*.

Bayley, N. (2006). *Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III). Technical Manual*. https://doi.org/10.1037/t14978-000

Bayley N, A. G. (2019). *Bayley Scales of Infant and Toddler Development-Fourth Edition Administration Manual* (fourth). NCS Pearson;

Bedford, H., Walton, S., & Ahn, J. (2013). Measures of child development: a review. *Policy Research Unit in the Health of Children, Young People and Families*, *June*, 1–104.

Berk, L. E. (2020). *Child Development* (9th ed., Vol. 21, Issue 1). Pearson.

Berls, A. T., & McEwen, I. R. (1999). Battelle developmental inventory. *Physical Therapy*, *79*(8), 776–783. https://doi.org/10.1093/ptj/79.8.776

Bian, X., Xie, H., Squires, J., & Chen, C. (2017). Adapting a parent-completed, socioemotional questionnaire in china: The ages & stages questionnaires: Social-emotional. *Infant Mental Health Journal*, *38*(2), 258–266.

Bian, X., Yao, G., Squires, J., Hoselton, R., Chen, C. I., Murphy, K., Wei, M., & Fang, B. (2012). Translation and use of parent-completed developmental screening test in Shanghai. *Journal of Early Childhood Research*, *10*(2), 162–175. https://doi.org/10.1177/1476718X11430071

Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the Differential Ability Scales in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, *116*(5), 331–343. https://doi.org/10.1352/1944-7558-116.5.331

Bliss, S. L. (2007). Test Reviews: Newborg, J.(2005). Battelle Developmental Inventory– Second Edition. Itasca, IL: Riverside. *Journal of Psychoeducational Assessment*, *25*(4), 409–415.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*.

Brewer, N., Young, R. L., & Lucas, C. A. (2020). Autism screening in early childhood: Discriminating autism from other developmental concerns. *Frontiers in Neurology*, *11*, 594381.

Brigance® Early Childhood Screens III. (n.d.). Curriculum Associates. https://www.curriculumassociates.com/products/brigance-screens-iii/brigance-early-childhood-screens-iii

Brigance, A. H. (2016). *Brigance early childhood screens III*. Curriculum Associates.

Brittany A. Dale, J. M. C., & Kaufman, Emily L. Winter, A. S. (n.d.). *Bayley-4 performance of very young children with autism and developmental delay*. https://doi.org/DOI: 10.1002/pits.22682

Brothers, K. B., Page Glascoe, F., & Robertshaw, N. S. (2008). PEDS: Developmental milestones - An accurate brief tool for surveillance and screening. *Clinical Pediatrics*, *47*(3), 271–279. https://doi.org/10.1177/0009922807309419

Burns, T. G., King, T. Z., & Spencer, K. S. (2013). Mullen scales of early learning: The utility in assessing children diagnosed with autism spectrum disorders, cerebral palsy, and epilepsy. *Applied Neuropsychology: Child*, *2*(1), 33–42. https://doi.org/10.1080/21622965.2012.682852

Bzoch, K. R., League, R., & Brown, V. L. (2003a). *Receptive-expressive Emergent Language Test: Examiner's Manual*. Pro-ed.

Bzoch, K. R., League, R., & Brown, V. L. (2003b). Receptive-expressive emergent language test third edition (REEL-3). *St. Antonio: Pearson*.

Carter, A. S., Briggs-Gowan, M. J., & Davis, N. O. (2004). Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *45*(1), 109–134. https://doi.org/10.1046/j.0021-9630.2003.00316.x

Charafeddine, L., Dani, A., Badr, L. K., Sinno, D., Tamim, H., Khoury, J., Nasser, F., & Makki, M. (2019). The psychometric properties of the Ages and Stages Questionnaires-3 in Arabic: Cross-sectional observational study. *Early Human Development*, *136*(May), 33–38. https://doi.org/10.1016/j.earlhumdev.2019.06.010

Chen, C.-Y., Xie, H., Filgueiras, A., Squires, J., Anunciação, L., & Landeira-Fernandez, J. (2017). Examining the psychometric properties of the Brazilian Ages & Stages Questionnaires-Social-Emotional: Use in public child daycare centers in Brazil. *Journal of Child and Family Studies*, *26*, 2412–2425.

Chunsuwan, I., Hansakunachai, T., & Pornsamrit, S. (2016). Parent Evaluation of Developmental Status (PEDS) in screening: The Thai experience. *Pediatrics International*, *58*(12), 1277–1283. https://doi.org/10.1111/ped.13055

Clark, H., Coll-Seck, A. M., Banerjee, A., Peterson, S., Dalglish, S. L., Ameratunga, S., Balabanova, D., Bhan, M. K., Bhutta, Z. A., & Borrazzo, J. (2020). A future for the world's children? A WHO–UNICEF–Lancet Commission. *The Lancet*, *395*(10224), 605–658.

Coghlan, D., Kiing, J. S. H., & Wake, M. (2003). Parents' evaluation of developmental status in the Australian day-care setting: Developmental concerns of parents and carers. *Journal of Paediatrics and Child Health*, *39*(1), 49–54. https://doi.org/10.1046/j.1440-1754.2003.00084.x

Colbert, A. M., Lamb, M. M., Asturias, E. J., Muñoz, F. M., Bauer, D., Arroyave, P., Hernández, S., Martínez, M. A., Paniagua-Avila, A., & Olson, D. (2020). Reliability and Validity of an Adapted and Translated Version of the Mullen Scales of Early Learning (AT-MSEL) in Rural Guatemala. *Child: Care, Health and Development*, *46*(3), 327–335.

Connery, A. K., Colbert, A. M., Lamb, M. M., Hernández, S., Martínez, M. A., Bauer, D., Arroyave, P., El Sahly, H. M., Paniagua-Avila, A., & Calvimontes, M. (2019). Receptive language skills among young children in rural Guatemala: The relationship between the Test de Vocabulario en Imagenes Peabody and a translated and adapted version of the Mullen Scales of Early Learning. *Child: Care, Health and Development*, *45*(5), 702–708.

Copeland, L. (2017). *Identifying and Addressing Developmental–Behavioral Problems: A Practical Guide for Medical and Nonmedical Professionals, Trainees, Researchers, and Advocates*. LWW.

Cunha, A. C. B., Berkovits, M. D., & Albuquerque, K. A. (2018a). Developmental assessment with young children: A systematic review of Battelle studies. *Infants & Young Children*, *31*(1), 69–90.

Cunha, A. C. B., Berkovits, M. D., & Albuquerque, K. A. (2018b). Developmental Assessment with Young Children: A Systematic Review of Battelle Studies. *Infants and Young Children*, *31*(1), 69–90. https://doi.org/10.1097/IYC.0000000000000106

Deeks, J. J., Wisniewski, S., & Davenport, C. (2013). Chapter 4: guide to the contents of a Cochrane Diagnostic Test Accuracy Protocol. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version*, *1*(0), 1–15.

Del Rosario, C., Slevin, M., Molloy, E. J., Quigley, J., & Nixon, E. (2021). How to use the Bayley Scales of Infant and Toddler Development. *Archives of Disease in Childhood: Education and Practice Edition*, *106*(2), 108–112. https://doi.org/10.1136/archdischild-2020-319063

Delehanty, A. D., Stronach, S., Guthrie, W., Slate, E., & Wetherby, A. M. (2018). Verbal and nonverbal outcomes of toddlers with and without autism spectrum disorder, language delay, and global developmental delay. *Autism & Developmental Language Impairments*, *3*, 2396941518764764.

Developmental Profile 4 Manual (Vol. 4, pp. 1–5). (2020).

DiStefano, C., & Kasari, C. (2016). The window to language is still open: Distinguishing between preverbal and minimally verbal children with ASD. *Perspectives of the ASHA Special Interest Groups*, *1*(1), 4–11.

Dockrell, J., Llaurado, A., Hurry, J., Cowan, R., & Flouri, E. (2017). Review of assessment measures in the early years. *Education Edowment Foundation*, *December*, 68. https://doi.org/10.13140/RG.2.2.33853.97765

Dumont, R., Willis, J. O., Viezel, K., & Zibulsky, J. (2014). Mullen Scales of Early Learning, AGS Edition, 1995. In *Encyclopedia of Special Education*. https://doi.org/10.1002/9781118660584.ese1602

Encyclopedia of Child Behavior and Development. (2011). *Reference Reviews*, *25*(6), 10–11. https://doi.org/10.1108/09504121111155932

Farmer, C., Adedipe, D., Bal, V., Chlebowski, C., & Thurm, A. (2019). *Reliability of the Vineland Adaptive Behavior Scales, Third Edition*. *301*.

Fernald, L. C. H., Prado, E., Kariger, P., & Raikes, A. (2017). *A toolkit for measuring early childhood development in low and middle-income countries*.

Flax, J. F., Realpe-Bonilla, T., Roesler, C., Choudhury, N., & Benasich, A. (2009). Using early standardized language measures to predict later language and early reading outcomes in children at high risk for language-learning impairments. *Journal of Learning Disabilities*, *42*(1), 61–75.

Floyd, R. G., Shands, E. I., Alfonso, V. C., Phillips, J. F., Autry, B. K., Mosteller, J. A., Skinner, M., & Irby, S. (2015). A Systematic Review and Psychometric Evaluation of Adaptive Behavior Scales and Recommendations for Practice. *Journal of Applied School Psychology*, *31*(1), 83–113. https://doi.org/10.1080/15377903.2014.979384

Frankenburg, W. K., Dodds, J., Archer, P., Shapiro, H., & Bresnick, B. (1992). The Denver II: A Major Revision and Restandardization of the Denver Developmental Screening Test. *Pediatrics*, *89*(1), 91–97. https://doi.org/10.1542/peds.89.1.91

Frankenburg, W. K., & Dodds, J. B. (1967). The Denver Developmental Screening Test. *The Journal of Pediatrics*, *71*(2), 181–191. https://doi.org/10.1016/S0022-3476(67)80070-2

French, B. F. (2013). *Brigance Screens III Technical Manual*.

Gannon, M. H. (2007). *VALIDITY OF THE DEVELOPMENTAL CHECKLIST OF THE DEVELOPMENTAL OBSERVATION CHECKLIST SYSTEM* (Issue August). https://core.ac.uk/download/pdf/43618068.pdf

Glascoe, F. P. (1997). Do the Brigance Screens Detect Developmental and Academic Problems? *Diagnostique*, *22*(2), 87–99. https://doi.org/10.1177/073724779702200202

Glascoe, F. P. (2000). Evidence-based approach to developmental and behavioural surveillance using parents concerns. *Child: Care, Health and Development*, 137–149.

Glascoe, F. P. (2002). The Brigance Infant and Toddler Screen: Standardization and validation. *Journal of Developmental and Behavioral Pediatrics*, *23*(3), 145–150. https://doi.org/10.1097/00004703-200206000-00003

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*(3), 173–179. https://doi.org/10.1002/mrdd.20068

Glascoe, F. P., Byrne, K. E., Ashford, L. G., Johnson, K. L., Chang, B., & Strickland, B. (1992). Accuracy of the Denver-II in developmental screening. *Pediatrics*, *89*(6), 1221–1225.

Glascoe, F. P., Marks, K. P., & Squires, J. (2012). Improving the Definition of Developmental Delay. *Journal of Developmental & Behavioral Pediatrics*, *33*(1). https://journals.lww.com/jrnldbp/Fulltext/2012/01000/Improving_the_Definition_of_Developmental_Delay.15.aspx

Hutchins, T. (2021). Rossetti Infant-Toddler Language Scale. In *Encyclopedia of Autism Spectrum Disorders* (pp. 4019–4022). https://doi.org/10.1007/978-3-319-91280-6_803

Janus, M., & Offord, D. R. (2007). Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, *39*(1), 1.

Kaiser, A. P., & Roberts, M. Y. (2011). Advances in early communication and language intervention. *Journal of Early Intervention*, *33*(4), 298–309.

Kanne, S. M., Gerber, A. J., Quirmbach, L. M., Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2011). The role of adaptive behavior in autism spectrum disorders: Implications for functional outcome. *Journal of Autism and Developmental Disorders*, *41*(8), 1007–1018. https://doi.org/10.1007/s10803-010-1126-4

Keen, D., Webster, A., & Ridley, G. (2016). How well are children with autism spectrum disorder doing academically at school? An overview of the literature. *Autism*, *20*(3), 276–294.

Kilburn, J. (2020). *Tables of Scores*.

Koshy, B., Srinivasan, M., Scharf, R., Strand, T. A., Mohan, V. R., Beulah, R., John, S., Muliyil, J., & Kang, G. (2024). Stability and predictability of Bayley Scales of Infant and Toddler Development: evidence from a south Indian birth cohort prospective study. *BMJ Open*, *14*(11), e082624.

Laws, G., & Hall, A. (2014). Early hearing loss and language abilities in children with Down syndrome. *International Journal of Language & Communication Disorders*, *49*(3), 333–342.

Lee Zimmerman, I., Steiner, V., & Evatt Pond, R. (2012). *Preschool Language Scales- Fifth Edition (PLS-5)*. 1–19.

Letts, E., King-Dowling, S., Calotti, R., DiCristofaro, N., & Obeid, J. (2023). Investigating the validity of the Ages and Stages Questionnaire to detect gross motor delays in a community sample of toddlers: A cross-sectional study. *Early Human Development*, *187*, 105882.

Levey, S. (2019). *Introduction to Language Development*. Plural Publishing.

Luiz, D. M., Foxcroft, C. D., & Tukulu, A. N. (2004). The Denver II Scales and the Griffiths Scales of Mental Development: a correlational study. *Journal of Child & Adolescent Mental Health*, *16*(2), 77–81. https://doi.org/10.2989/17280580409486573

Lyons, M. (2021). Preschool Language Scale–5. In *Encyclopedia of Autism Spectrum Disorders* (pp. 3658–3659). Springer.

Macy, M., Bagnato, S. J., Macy, R. S., & Salaway, J. (2015). Conventional tests and testing for early intervention eligibility: Is there an evidence base? *Infants & Young Children*, *28*(2), 182–204.

Mantzicopoulos, P. (1999). Reliability and validity estimates of the brigance K&1 screen based on a sample of disadvantaged preschoolers. *Psychology in the Schools*, *36*(1), 11–19. https://doi.org/10.1002/(sici)1520-6807(199901)36:1<11::aid-pits2>3.0.co;2-r

Martin, S. S., & Whiteman, V. (2017). The use of the Brigance Early Childhood Screens III with preschool-aged children. *Journal of Psychoeducational Assessment*, *35*(6), 584–596. https://doi.org/10.1177/0734282916679852

Martinussen, M., & Valla, L. (2013). *Måleegenskaper ved den norske versjonen av Ages and Stages Questionnaires (ASQ)*.

McCoy, D. C., Waldman, M., Team, C. F., & Fink, G. (2018). Measuring early childhood development at a global scale: Evidence from the Caregiver-Reported Early Development Instruments. *Early Childhood Research Quarterly*, *45*, 58–68.

McLeod, S., Crowe, K., Masso, S., Baker, E., McCormack, J., Wren, Y., Roulstone, S., & Howland, C. (2017). Profile of Australian preschool children with speech sound disorders at risk for literacy difficulties. *Australian Journal of Learning Difficulties*, *22*(1), 15–33.

Meisels, S. J., Wen, X., & Beachy-Quick, K. (2010a). Authentic assessment for infants and toddlers: Exploring the reliability and validity of the ounce scale. *Applied Developmental Science*, *14*(2), 55–71. https://doi.org/10.1080/10888691003697911

Meisels, S. J., Wen, X., & Beachy-Quick, K. (2010b). Authentic assessment for infants and toddlers: Exploring the reliability and validity of the Ounce Scale. *Applied Developmental Science*, *14*(2), 55–71.

Milne, S., McDonald, J., & Comino, E. J. (2012). The use of the bayley scales of infant and toddler development III with clinical populations: A preliminary exploration. *Physical and Occupational Therapy in Pediatrics*, *32*(1), 24–33. https://doi.org/10.3109/01942638.2011.592572

Milosavljevic, B., Vellekoop, P., Maris, H., Halliday, D., Drammeh, S., Sanyang, L., Darboe, M. K., Elwell, C., Moore, S. E., & Lloyd-Fox, S. (2019). Adaptation of the Mullen Scales of Early Learning for use among infants aged 5- to 24-months in rural Gambia. *Developmental Science*, *22*(5), 1–17. https://doi.org/10.1111/desc.12808

Mokkink, L. B., De Vet, H. C. W., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*, 1171–1179.

Moon, K. (2019). *The Reliability of Early Language Assessment Tools*. https://vtext.valdosta.edu/xmlui/bitstream/handle/10428/3934/moon-yae-eun_thesis_2019.pdf?sequence=1&isAllowed=y

Mukherjee, S. B., Verma, S., Sharma, S., & Aneja, S. (2022). Diagnostic Accuracy of Parents' Evaluation of Developmental Status (PEDS), PEDS Developmental Milestones, and PEDS Combined in Indian Children Aged Less than 2 Years. *Indian Journal of Pediatrics*, *89*(5), 459–465. https://doi.org/10.1007/s12098-020-03651-y

Mullen, E. M. (1995). *Mullen scales of early learning*. AGS Circle Pines, MN.

Newborg, J. (2020). *Battelle Developmental Inventory, Third Edition (BDI-3)*.

Newborg, J., & Company, R. P. (2005). *Battelle developmental inventory*. Riverside Pub.

Nimkar, S., Joshi, S., Kinikar, A., Valvi, C., Devaleenal, D. B., Thakur, K., Bendre, M., Khwaja, S., Ithape, M., Kattagoni, K., Paradkar, M., Gupte, N., Gupta, A., Suryavanshi, N., Mave, V., Dooley, K. E., & Arenivas, A. (2021). Mullen Scales of Early Learning Adaptation for Assessment of Indian Children and Application to Tuberculous Meningitis. *Journal of Tropical Pediatrics*, *67*(3), 1–9. https://doi.org/10.1093/tropej/fmaa034

Parmeggiani, A., Corinaldesi, A., & Posar, A. (2019). Early features of autism spectrum disorder: a cross-sectional study. *Italian Journal of Pediatrics*, *45*, 1–8.

Paul, R. (2007). *Language disorders from infancy through adolescence: Assessment & intervention* (Vol. 324). Elsevier Health Sciences.

Pepperdine, C. R., & McCrimmon, A. W. (2018). Test Review: Vineland Adaptive Behavior Scales, Third Edition (Vineland-3) by Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. . *Canadian Journal of School Psychology*, *33*(2), 157–163. https://doi.org/10.1177/0829573517733845

Piñon, M. (2010). Theoretical Background and Structure of the Bayley Scales of Infant and Toddler Development, Third Edition. *Bayley-III Clinical Use and Interpretation*, 1–28.

https://doi.org/10.1016/B978-0-12-374177-6.10001-7

Pyle, N., Simpson, D., & Vanderwood, M. (2018). Measuring child outcomes with the Brigance Early Childhood Screens III. *Journal of Early Intervention*, *40*(2), 137–152. https://doi.org/10.1177/1053815118770847

Restrepo, M. A., & Silverman, S. W. (2001). Assessing the Validity of Language Measures. *American Journal of Speech-Language Pathology* , *10*(November), 382–393. https://doi.org/10.1044/1058-0360(2001/032)CITATIONS

Reuter, J., Katoff, L., & Wozniak, J. R. (2000). *Kent inventory of developmental skills*. Los Angeles: Western Psychological Services.

Riley, E., Paynter, J., & Gilmore, L. (2019). Comparing the Mullen Scales of Early Learning and the Preschool Language Scale—Fifth Edition for Young Children with Autism Spectrum Disorder. *Advances in Neurodevelopmental Disorders*, *3*(1), 29–37. https://doi.org/10.1007/s41252-018-0084-2

Ringwalt, S. (2008a). Developmental Screening and Assessment Instruments with an Emphasis on Social and Emotional Development for Young Children Ages Birth through Five. *National Early Childhood Technical Assistance Center …*, *May*. http://eric.ed.gov/?id=ED505971

Ringwalt, S. (2008b). Developmental screening and assessment instruments with an emphasis on social and emotional development for young children ages birth through five. *British Medical Journal*, *2*. https://doi.org/10.1136/bmj.2.5962.87

Ritvo, A. R. (2013). Encyclopedia of Autism Spectrum Disorders. In *Encyclopedia of Autism Spectrum Disorders*. https://doi.org/10.1007/978-1-4419-1698-3

Rossetti, L. M. (2006). *The Rossetti infant-toddler language scale*. LinguiSystems East Moline, IL.

Rousseau, M., Dionne, C., Savard, R. T., Schonhaut, L., & Londono, M. (2021). Translation and Cultural Adaptation of the Ages and Stages Questionnaires (ASQ) Worldwide: A Scoping Review. *Journal of Developmental and Behavioral Pediatrics : JDBP*, *42*(6), 490–501. https://doi.org/10.1097/DBP.0000000000000940

Sahli, A. S., & Belgin, E. (2017). Adaptation, validity, and reliability of the Preschool Language Scale–Fifth Edition (PLS–5) in the Turkish context: The Turkish Preschool Language Scale–5 (TPLS–5). *International Journal of Pediatric Otorhinolaryngology*, *98*, 143–149. https://doi.org/10.1016/j.ijporl.2017.05.003

Saldana, J. (2018). *Writing qualitatively: The selected works of Johnny Saldaña*. Routledge.

Santrock, J. . (2014). *Life Span Development* (9th ed., Vol. 13, Issue 1). Kevin Feyen.

Schludermann, S. (2001). Book Review: Adjustment of Adolescents: Cross-cultural similarities and differences. *International Journal of Behavioral Development*, *25*(2), 187–187. https://doi.org/10.1080/01650250042000023

Sheel, H., Suárez, L., & Marsh, N. V. (2023). Parents' Evaluation of Developmental Status and Strength and Difficulties Questionnaire as Screening Measures for Children in India: A Scoping Review. *Pediatric Reports*, *15*(1), 175–196. https://doi.org/10.3390/pediatric15010014

Singh, A., Yeh, C. J., & Boone Blanchard, S. (2017). Ages and Stages Questionnaire: a global screening scale. *Boletín Médico Del Hospital Infantil de México (English Edition)*, *74*(1), 5–12. https://doi.org/10.1016/j.bmhime.2016.07.001

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales, Third Edition (Vineland-3)* (3rd ed.). Pearson.

Sperhac, A. M., & Salzer, J. L. (1976). A new developmental screening test. The Denver II. *Journal of the American Academy of Nurse Practitioners*, *3*(5), 570–572.

Squires, J., & Bricker, D. (2009). *Ages and Stages Questionnaire (ASQ): A Parent Completed Child Monitoring System (3rd ed.)*.

Squires, J., Ph, D., Bricker, D., Twombly, E., Yockelson, S., Davis, M. S., & Kim, Y. (2002). A GES & S TAGES Q UESTIONNAIRES ®: S OCIAL -E MOTIONAL A P ARENT -C OMPLETED , C HILD -M ONITORING S YSTEM FOR S OCIAL -E MOTIONAL B EHAVIORS by About this CD-ROM. *Human Development*.

Squires, J., Twombly, E., Bricker, D., & Potter, L. (2009). ASQ-3 Technical Report. *ASQ-3™ User's Guide*. https://agesandstages.com/wp-content/uploads/2017/05/ASQ-3-Technical-Appendix_web.pdf

Stone-MacDonald, A., Pizzo, L., & Feldman, N. (2018). Fidelity of implementation in assessment of infants and toddlers: Evaluating developmental milestones and outcomes. *Fidelity of Implementation in Assessment of Infants and Toddlers: Evaluating Developmental Milestones and Outcomes*, 1–177. https://doi.org/10.1007/978-3-319-74618-0

Swartzmiller, M. D. (2014a). *Test review: Developmental assessment of young children–second edition (DAYC-2)*. SAGE Publications Sage CA: Los Angeles, CA.

Swartzmiller, M. D. (2014b). Test Review: Developmental Assessment of Young Children–Second Edition (DAYC-2). *Journal of Psychoeducational Assessment*, *32*(6), 577–580. https://doi.org/10.1177/0734282913518380

Swineford, L. B., Guthrie, W., & Thurm, A. (2015). Convergent and divergent validity of the mullen scales of early learning in young children with and without autism spectrum disorder. *Psychological Assessment*, *27*(4), 1364–1378. https://doi.org/10.1037/pas0000116

Tambyraja, S. R., Schmitt, M. B., Farquharson, K., & Justice, L. M. (2017). Home literacy environment profiles of children with language impairment: associations with caregiver-and child-specific factors. *International Journal of Language & Communication Disorders*, *52*(2), 238–249.

Technical, E. A. (2021). Receptive-Expressive Emergent Language Test (REEL-3). *Encyclopedia of Autism Spectrum Disorders*, 3886–3888.

Thomson Delmar Learning. (2007). *CHILDREN ' S DEVELOPMENT: Denver Screening*.

Vameghi, R., Sajedi, F., Shahshahani, S., & Biglarian, A. (2015). Validity and reliability determination of Parents Evaluation of Developmental Status (PEDS) in 4-60 months old children in Tehran. *Iranian Rehabilitation Journal*, *13*(2), 62–67.

Volkmar, F. R., Rogers, S. J., Paul, R., & Pelphrey, K. A. (2014). *Handbook of autism and pervasive developmental disorders, volume 1: diagnosis, development, and brain mechanisms*. John Wiley & Sons.

Voress, J. K., & Maddox, T. (2013). *Developmental assessment of young children: Examiner's manual*. Pro-Ed.

Ward, C., Chiat, S., & Townend, G. S. (2021). A comparison of formal and informal methods for assessing language and cognition in children with Rett syndrome. *Research in Developmental Disabilities*, *114*, 103961.

Wiig, E. H., Semel, E., & Secord, W. A. (2014). Which test is right for the youngest children on your caseload ? *Pearson Clinical Assessment*.

Woolfenden, S., Eapen, V., Williams, K., Hayen, A., Spencer, N., & Kemp, L. (2014). A systematic review of the prevalence of parental concerns measured by the Parents' Evaluation of Developmental Status (PEDS) indicating developmental risk. *BMC Pediatrics*, *14*(1), 1–13. https://doi.org/10.1186/1471-2431-14-231

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1979). *Preschool language scale*.