

Additional file 2. Assessment Instrument Descriptions and Scoring

Ages and Stages Questionnaire (ASQ-3 & ASQ:SE-2)(1,2)	
Domains and Scoring	The ASQ-3 evaluates 5 domains of development: communication, gross motor, fine motor, problem solving, and personal-social. Each domain has a set of 6 items and parents rate the most appropriate answer for the presence of each skill: “Yes,” “Sometimes,” “Not Yet,” with point values of 10, 5, or 0, respectively. Each domain question set is totalled independently. Prespecified clinical cut-points for each subscale are provided to indicate whether the score falls within a normal developmental range based on chronological age, or if it represents “at risk” or delayed development(2)
Brief Infant-Toddler Social Emotional Assessment (BITSEA)(3)	
Domains and Scoring	BITSEA: The BITSEA has items that measure certain areas such as externalizing problems, internalizing problems, dysregulation, maladaptive behaviors and atypical behaviors as well as competence. These areas combine to form two scales on the measure: the Problems scale and the Competencies scale. According to the examiner’s manual, on Problems scales, a percentile rank of 25 or higher are termed “possible problem”. Similarly, percentile rank of 15 or lower on the Competence scale are termed “possible deficit/delay range”.
Child Behavior Checklist (CBCL)(4)	
Domains and Scoring	The CBCL/1.5-5 obtains caregivers' ratings of 99 problem items. Items are scored on the following syndrome scales: Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Attention Problems, Aggressive Behavior, and Sleep Problems. Items are also scored on the following DSM-oriented scales: Affective Problems, Anxiety Problems, Pervasive Developmental Problems, Attention Deficit/Hyperactivity Problems, Stress Problems, Autism Spectrum Problems, and Oppositional Defiant Problems. The measure also includes open-ended questions to obtain additional qualitative information. Problem-related questions allow the respondent to write in additional problems that were not listed in the 99 items and to further describe problems, physical and mental disabilities, and primary concerns about the child. The strengths-based question allows the respondent to describe the best things about the child. The CBCL is to be completed by the parent/caretaker who spends the most time with the child.
Childhood Executive Functioning Inventory (CHEXI)(5)	
Domains and Scoring	Subscale 1: Working memory Mean score for the following items: 1, 3, 6, 7, 9, 19, 21, 23, 24 Subscale 2: Planning Mean score for the following items: 12, 14, 17, 20 Subscale 3: Regulation Mean score for the following items: 2, 4, 8, 11, 15 Subscale 4: Inhibition Mean score for the following items: 5, 10, 13, 16, 18, 22
Griffiths III Mental Development Scales(6)	
Domains and Scoring	The Scales measure development trends which are significant for intelligence, or indicative of functional mental growth in babies and young children. In the GMDS-ER, development was measured up to 8 years of age. The new Griffiths III measures maturation up to 6 years.

	Griffiths III represents an evolution in developmental assessment, based on the latest research and theory, and building on the best of the past. Griffiths-III is a child-friendly, play-oriented measure for assessing the rate of development in infants from birth to 6 years. Coral Group et al used adapted questions from the Griffiths; 10 skills were assessed: crawl, side step along furniture, stand alone, pick up tiny objects with thumb and index finger (pincer grip), put one brick upon another (stack bricks), finger feed, know their own name, has one definite and meaningful word, point at objects and wave 'bye-bye'. Questions were based on the parental report section of the Griffiths Scales of Mental Development Version 2, 0–2 Years
Heads–Toes–Knees–Shoulders Task (HTKS)(7)	
Domains and Scoring	Children are asked to play a game in which they must do the opposite of what the experimenter says. The experimenter instructs children to touch their head (or their toes), but instead of following the command, the children are supposed to do the opposite and touch their toes. If children pass the head/toes part of the task, they complete an advanced trial where the knees and shoulders commands are added. The HTKS task has been conceptualized by Ponitz, et al., (2008) as a measure of inhibitory control (a child must inhibit the dominant response of imitating the examiner), working memory (a child must remember the rules of the task) and attention focusing (must focus attention to the directions being presented by the examiner)
Kinder Infant Development Scale (KIDS)(8)	
Domains and Scoring	The KIDS consists of 9 subscales for behavior: “physical motor,” “manipulation,” “receptive language,” “expressive language,” “language concepts,” “social relationships with children,” “social relationships with adults,” “discipline,” and “feeding”. In each subscale, behaviors that can be performed by the child are checked. Using the 9 areas of development, this test produces a clear profile of the developmental age and developmental quotient of the child.
Korean Developmental Screening Test (K-DST)(9)	
Domains and Scoring	K-DST is composed of six domains (gross motor, fine motor, cognition, language, sociality, and self-care) and each domain consists of eight questions. The selfcare category is only applicable to participants aged 18 months or older. Each question was scored on a scale of 0–3 and each domain had a maximum attainable score of 24 points
LENA Developmental Snapshot (LDS)(10)	
Domains and Scoring	LDS: uses parents' reports to assess children's expressive vocabularies & word combinations, as well as risk factors for language delays. The scores represent a correlation based on standardized testing, that indicates the risk level for delay in each respective domain.
Modified Checklist for Autism in Toddlers, Revised (MCHAT–R)(11)	
Domains and Scoring	Neurodevelopment Assessment Screening Tool for Autism Total Score 0-2: The score is LOW risk. No Follow-Up needed. Child has screened negative. Rescreen at 24 months if the child is younger than 2 years old (or after 3 months has elapsed), and refer as needed if developmental surveillance or other tools suggest risk for ASD. Total Score 3-7 : The score is MODERATE risk. Administer the M-CHAT-R Follow-Up items that correspond to the at-risk responses. Only those items which were scored at risk need to be completed. If 2 or more items continue to be at-risk, refer the child immediately for (a) early intervention and (b) diagnostic evaluation.

	Total Score: 8-20 : The score is HIGH risk. It is not necessary to complete the M-CHAT-R Follow-Up at this time. Bypass Follow-Up, and refer immediately for (a) early intervention and (b) diagnostic evaluation.
NIH Toolbox Early Childhood Cognition Battery(12)	
Domains and Scoring	The NIH Toolbox Early Childhood Cognition Battery is recommended for ages 4-6. It produces individual test scores and an Early Childhood Composite Score (includes Dimensional Change Card Sort, Flanker, Picture Sequence Memory (Form A), Picture Vocabulary, and Speeded Matching). Other Cognition tests can be administered to understand the participant's functioning, but are not required for the calculation of the composite scores.
Survey of Well-Being of Young Children (SWYC) Developmental Milestones(13)	
Domains and Scoring	The SWYC assesses multiple domains of children's well-being; the SWYC Milestones assess the child's cognitive, language, and motor development. Each form includes 10 items. 1.Score each item using these values: Not yet=0; Somewhat=1, Very Much=2. 2. Add up all 10 item scores to calculate the total score. If answers are missing, reevaluate if possible. Otherwise, missing items count as zero. 3. On the Milestones Scoring Chart, the child's age in months is indicated in the far left column. Check to be sure that the parent completed the correct form for the child's age (second column). If not, the score will be misleading. 4. Following along the appropriate age row, try to find the number corresponding to the child's total score. If you find the appropriate number, circle it.
Questionnaire adapted from the national well child assessment by public health nurses(14)	
Domains and Scoring	The assessment consisted of the following: (1) Can your child say three or more meaningful words such as 'wan wan' (meaning a dog in Japanese)? (2) Does your child follow simple commands such as 'Please throw away the trash'? (3) Does your child point to objects that they know, such as when asked to point to a 'wan wan'? As for question (3), public health nurses further check whether the child can perform pointing using picture cards. If the child does not satisfactorily meet the conditions represented by any of three questions, then the public health nurse considers that the child requires follow-up for language development. We defined the secondary outcome as the proportion of children who cannot say three or more meaningful words (the first question). These three questions were based on the guidance for health check-ups in infants and children provided by the Ministry of Health, Labour and Welfare and the National Center for Child Health and Development in Japan.
Strengths and Difficulties Questionnaire (SDQ)(15)	
Domains and Scoring	All versions of the SDQ ask about 25 attributes, some positive and others negative. These 25 items are divided between 5 scales: 1) emotional problems (5 items) } 1) to 4) added together to generate a total difficulties score (based on 20 items) 2) conduct problems (5 items) 3) hyperactivity/inattention (5 items) 4) peer relationship problems (5 items) 5) prosocial behaviour (5 items)

1. Agarwal PK, Xie H, Sathyapalan Rema AS, Rajadurai VS, Lim SB, Meaney M, et al. Evaluation of the Ages and Stages Questionnaire (ASQ 3) as a developmental screener at 9, 18, and 24 months. *Early Hum Dev.* 2020 Aug 1;147:105081.
2. Quigg TC, Mahajerin A, Sullivan PD, Pradhan K, Bauer NS. Ages and Stages Questionnaires-3 Developmental Screening of Infants and Young Children With Cancer. <http://dx.doi.org/10.1177/1043454213493510> [Internet]. 2013 Jul 18 [cited 2024 Jul 1];30(5):235–41. Available from: <https://journals.sagepub.com/doi/abs/10.1177/1043454213493510>
3. Briggs-Gowan MJ, Carter AS, Irwin JR, Wachtel K, Cicchetti D V. The Brief Infant-Toddler Social and Emotional Assessment: Screening for Social-Emotional Problems and Delays in Competence. *J Pediatr Psychol* [Internet]. 2004 Mar 1 [cited 2024 Dec 15];29(2):143–55. Available from: <https://dx.doi.org/10.1093/jpepsy/jsh017>
4. for Addiction C, Health M. Child Behavior Checklist (CBCL) Brief description [Internet]. Available from: www.aseba.org
5. Thorell LB, Nyberg L. The Childhood Executive Functioning Inventory (CHEXI): A new rating instrument for parents and teachers. *Dev Neuropsychol.* 2008 Jul;33(4):536–52.
6. Griffiths R. The abilities of young children : a comprehensive system of mental measurement for the first eight years of life. (No Title) [Internet]. [cited 2024 Dec 15]; Available from: <https://cir.nii.ac.jp/crid/1130282272138548096>
7. Cameron Ponitz CE, McClelland MM, Jewkes AM, Connor CM, Farris CL, Morrison FJ. Author's personal copy Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Child Res Q.* 2008;23:141–58.
8. Miyake K, Ohmura M, Takashima M, Yamanouchi S, Hashimoto... - Google Scholar [Internet]. [cited 2024 May 13]. Available from: https://scholar.google.ca/scholar?hl=en&as_sdt=0%2C5&q=Miyake+K%2C+Ohmura+M%2C+Takashima+M%2C+Yamanouchi+S%2C+Hashimoto+T.+Kinder+Infant+Development+Scale+%28KIDS%29.+Center+of+Developmental+Education+and+Research%3B+1989.+&btnG=
9. Kim D, Choe YJ, Durrani BAZ, Kim E, Byeon J, Eun BL. Korean Developmental Screening Test for Infants and Children (K-DST): development, applications, and implications for future early childhood development interventions. Vol. 66, *Clinical and Experimental Pediatrics*. Korean Pediatric Society; 2023. p. 288–93.
10. Developmental Snapshot [Internet]. [cited 2024 Dec 15]. Available from: <https://www.lena.org/developmental-snapshot/>
11. M-CHAT™ - Autism Screening [Internet]. [cited 2024 Dec 15]. Available from: <https://www.mchatscreen.com/>
12. Zelazo PD, Anderson JE, Richler J, Wallner-Allen K, Beaumont JL, Weintraub S. II. NIH TOOLBOX COGNITION BATTERY (CB): MEASURING EXECUTIVE FUNCTION AND ATTENTION. *Monogr Soc Res Child Dev* [Internet]. 2013 Aug 1 [cited 2024 Dec 17];78(4):16–33. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/mono.12032>
13. C. Weitzman, L. Wegner, The Section on Developmental... - Google Scholar [Internet]. [cited 2024 May 13]. Available from: https://scholar.google.ca/scholar?hl=en&as_sdt=0%2C5&q=C.+Weitzman%2C+L.+Wegn

er%2C+The+Section+on+Developmental+and+Behavioral+Pediatrics%2C+Committee+o
n+Psychosocial+Aspects+of+Child+and+Family+Health%2C+Council+on+Early+Childho
od%2C+and+Society+for+Developmental+and+Behavioral+Pediatrics%2C+N.J.+Blum%
2C+M.M.+Macias%2C+N.S.+Bauer%2C+et+al.%2C+Promoting+optimal+development%
3A+screening+for+behavioral+and+emotional+problems%2C+Pediatrics+135+%282%29
+%282015+Feb+1%29+384%E2%80%933395.++&btnG=

14. National Center for Child Health and Development. Guidance for health checkups of infants and children.
15. Goodman R. The strengths and difficulties questionnaire: A research note. J Child Psychol Psychiatry. 1997 Jul;38(5):581–6.