Introducing NIH Baby Toolbox®

Assessment of

Cognition **Motor** **Social-Emotional**

Functioning

Available in Spring 2025

Download on the App Store

Ages 1-42 months

NIHBabyToolbox.org

NIH Baby Toolbox

Contains more than 30 assessments of Cognition, Motor, and Social-Emotional Functioning domains in one iPad app. Our valid, reliable, and norm-referenced measures allow for assessment of children throughout infancy and early childhood (ages 1 to 42 months).

Development At-A-Glance

2019

The National Institute of Child Health and Human Development awarded the NIH Baby Toolbox contract to Northwestern University using NIH Blueprint funds. Dr. Richard Gershon, Principal Investigator, and a team of 48 researchers started work to develop a brief, standardized assessment of neuropsychological, cognitive and social assessment of infants and toddlers ages 1-42 months.

2015

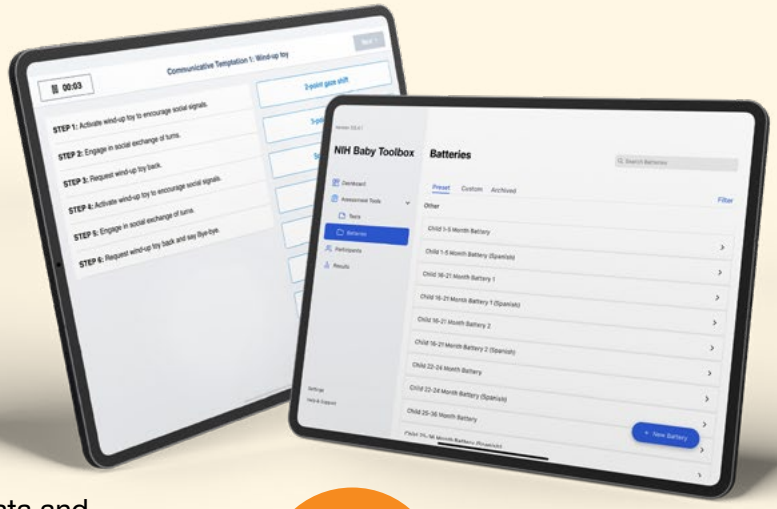
In response to demand for a more portable assessment system, Version 2 of the NIH Toolbox launched to researchers and clinicians on the App Store for iPad.

20

NIH Baby Toolbox be

Key Features in the NIH Baby Toolbox

- MODERN UX/UI CAPABILITIES CONSISTENT WITH THE TOOLBOX SYSTEMS
- USER-FRIENDLY ADMINISTRATION
- GAZE-BASED PREFERENTIAL LOOKING TASKS
- NORMATIVE SCORES FROM 1-42 MONTHS
- MULTI-METHOD ASSESSMENT MODALITIES



21

egan early validation.

2023

Collected normative data and validity evidence for the NIH Baby Toolbox tests aligned with the U.S. Census.

2024

The analytics team for the NIH Baby Toolbox developed scores and norms for the iPad-based tests.

2025

Public release is expected in late February.

Overview of Development

The NIH Baby Toolbox app assesses key developmental domains. Candidate measure selection was informed by input from an expert survey of developmental clinicians and researchers and a scoping review of over 30,000 manuscripts. Potential tests were reviewed for ease of tablet-based administration and scoring, likelihood of standardization, scientific value, and ability to present stimuli in both English and Spanish. In this way, the Baby Toolbox tests were adapted from existing measures and laboratory-based paradigms common among developmental researchers.

Key to the NIH Baby Toolbox is an emphasis on modern measurement methods. These methodologies include utilization of gaze-based preferential looking paradigms, on-device video recording and playback, and reporting functionality. To ensure the success of these developments, all of the NIH Baby Toolbox measures went through multiple stages of pilot testing and refinement. This included feasibility testing of new technology (e.g., four rounds of gaze-based pilot tests), evaluation of convergent validity between candidate Baby Toolbox measures and existing developmental tests, as well as a comprehensive “dry run” of the entire battery.

In preparation for the NIH Baby Toolbox release, a comprehensive norming study was conducted. Over 2,500 infants and toddlers from English- and Spanish-speaking households between the ages of 1 and 48 months were recruited across the United States. Participant demographics were matched to the U.S. Census for language spoken at home, race, ethnicity, sex assigned at birth, parental education, and geographic census region. Concurrent with the norming study, test-retest reliability and convergent validity studies were also conducted.

The NIH Baby Toolbox app is programmed using a similar user interface as the NIH Toolbox Version 3, which incorporates input from usability and accessibility experts to enhance its user experience for both examiners and participants. Users will find an easy-to-navigate interface, allowing for easy selection and administration of tests. Once the assessment is complete, results can be accessed through a Score Report or .CSV file designed for easy importing into research and clinical databases.

Cognition refers to the mental processes involved in attention, executive functioning, gaining knowledge, and comprehension, such as thinking, knowing, remembering, judging, and problem solving.

Through gaze-based and touch-based tasks on the iPad, behavioral observation, and caregiver report, the NIH Baby Toolbox Cognition domain produces individual measure-level scores and composite scores. Supplemental tests are not used in the calculation of composites but can be administered to better understand the participant's cognitive functioning.

COGNITION

Cognition Tests

SUBDOMAIN	MEASURE	AGE IN MONTHS	DESCRIPTION	MODALITY	ADMIN TIME IN MINUTES
Language	Looking While Listening	6-24	Assesses object-labeling using a gaze-based preferential looking paradigm. Infants hear a prompt labeling one of two images presented on a screen.	Gaze	5
Language	Picture Vocabulary	≥25	Assesses word knowledge. Children are asked to select one of four images via a touch-based task.	Touch	3
Language	MacArthur-Bates Communicative Development Inventories (CDI) Computerized Adaptive Tests (CAT) Production	9-30	Assesses vocabulary production. Computerized Adaptive Test version of the classic MCDI parent report and requires responses to only 20-50 words.	Caregiver Report	2
Language	MacArthur-Bates Communicative Development Inventories (CDI) Computerized Adaptive Tests (CAT) Comprehension	7-18	Assesses vocabulary comprehension. Computerized Adaptive Test version of the classic MCDI parent report and requires responses to only 20-50 words.	Caregiver Report	2
Language	Mullen Receptive	1-42	Assesses receptive language and communication skills based on elicited prompts.	Observation	3
Language	Mullen Expressive	1-42	Assesses expressive language and communication skills based on elicited prompts and observations.	Observation	3
Executive Function/ Cognition	Mullen Visual Reception	1-42	Assesses visual reception using both elicited prompts and touch-based responses on the iPad.	Observation/ Touch (Age Dependent)	4
Executive Function/ Cognition	Familiarization	6-21	Assesses ability to habituate to a visual stimulus.	Gaze	4

Note: All NBT Tests require additional equipment and/or manipulatives to properly administer.

COGNITION

Cognition Tests (Continued)

SUBDOMAIN	MEASURE	AGE IN MONTHS	DESCRIPTION	MODALITY	ADMIN TIME IN MINUTES
Executive Function/ Cognition	Visual Delayed Response	22-42	Assesses attention, executive functioning, and memory via a visual-delayed touch-based paradigm (22-42 months).	Touch	3
Executive Function/ Cognition	Delayed Memory	22-42	A touch-based assessment of delayed memory. Children are asked to encode items from Memory Task Learning and provide recognition of individual items after a 6–8-minute delay during Memory Task Test.	Touch	4
Math	Who Has More	25-42	Assesses approximate number system by asking children to select which of two numerical quantities on a screen is larger.	Touch	3
Math	Subitizing	25-42	Assesses approximate number system and number recognition by asking children to report on the number of objects presented briefly on a screen.	Observation	1
Math	Verbal Arithmetic	37-42	Assesses addition and subtraction by asking children to respond to simple arithmetic questions verbally.	Observation	2
Math	Counting	25-42	Assesses counting and cardinality by asking children to count as high as possible and to report the number of objects presently on a screen.	Observation	2
Math	Numerical Change Detection (supplementary)	6-24	Uses gaze tracking and passive viewing to assess number sense, approximate number system, and relative quantities.	Gaze	4

Note: All NBT Tests require additional equipment and/or manipulatives to properly administer.

Social-Emotional Functioning refers to early social competencies, including social responsiveness, communication, and relationships, as well as the experience, expression, and management of any strong feelings (i.e., emotions).

It also includes the experience, expression, and management of any strong feelings (i.e., emotions), such as excitement, fear, or anger, which can be positive or negative. Early social-emotional functioning is foundational for children's health and development and a meaningful predictor of school readiness, academic achievement, social competence, and psychopathology.

SOCIAL-EMOTIONAL FUNCTIONING

Social-Emotional Tests

SUBDOMAIN	MEASURE	AGE IN MONTHS	DESCRIPTION	MODALITY	ADMIN TIME IN MINUTES
Temperament	Infant Behavior Questionnaire - Revised - Very Short Form (IBQ-R-VSF)	3-12	An assessment of infant's effortful control, surgency, negative affectivity, and fear of unfamiliar adults.	Caregiver Report	6
Temperament	Early Childhood Behavior Questionnaire - Very Short Form (ECBQ-VSF)	13-36	An assessment of young children's effortful control, surgency, and negative affectivity.	Caregiver Report	6
Temperament	Children's Behavior Questionnaire - Very Short Form (CBQ-VSF)	37-42	An assessment of children's effortful control, surgency, and negative affectivity.	Caregiver Report	6
Negative Affect	PROMIS EC Parent-Report Bank v1.0 - Anger/Irritability	13-42	An assessment of children's angry mood, including irritability and grouchiness, and angry behavior, such as tantrums.	Caregiver Report/ CAT	6
Negative Affect	PROMIS EC Parent-Report Bank v1.0 - Anxiety	13-42	An assessment of children's feelings of fear, anxious misery, hyperarousal, and social/separation anxiety.	Caregiver Report/ CAT	1
Negative Affect	PROMIS EC Parent-Report Bank v1.0 - Depressive Symptoms	13-42	An assessment of children's feelings of sadness and withdrawal, negative self-views, and anhedonia.	Caregiver Report/ CAT	1
Psychological Well-Being	PROMIS EC Parent-Report Bank v1.0 - Positive Affect	13-42	An assessment of children's mood and feelings associated with positive affective experiences, such as happiness, joy, and enthusiasm.	Caregiver Report/ CAT	1
Self-Regulation	PROMIS EC Parent-Report Scale v1.0 - Self-Reg - Flexibility 5a	13-42	An assessment of children's ability to adapt in response to environmental demands, changes, and expectations.	Caregiver Report	1
Self-Regulation	PROMIS EC Parent-Report Scale v1.0 - Self-Reg - Frustration Tolerance 6a	13-42	An assessment of children's ability to recognize and regulate emotions and behaviors to serve one's own goals.	Caregiver Report	1
Social Relationships	PROMIS EC Parent-Report SF v1.0 - Soc Rel - Child-Caregiver Interactions	13-42	An assessment of positive child-caregiver interactions, experiences, and connectedness.	Caregiver Report	1
Social Relationships	PROMIS EC Parent-Report SF v1.0 - Soc Rel - Peer Relationships 4a	13-42	An assessment of children's positive interactions with peers, sociability, and empathic behaviors.	Caregiver Report	1
Social Relationships, Social Communication	Caregiver Checklist	6-42	An assessment of children's interactions with people and the environment	Caregiver Report	2
Social Relationships, Social Communication	Social Observation	9-42	An assessment of children's social functioning, including social attention, social communication, play, and prosocial behavior in response to specific, standardized verbal prompts and probes, and during more spontaneous social interactions with the caregiver.	Observation	10

Note: All NBT Tests require additional equipment and/or manipulatives to properly administer.

Motor function involves complex physiological processes and requires the integration of multiple systems, including neuromuscular, musculoskeletal, cardiopulmonary, neural motor, and sensory-perceptual systems.

The NIH Baby Toolbox motor measures assess manual dexterity, manual planning, locomotion, postural transitions, and stationary postural control. The tests produce Fine Motor and Gross Motor scores as well as an overall Motor composite score. Supplemental measures are not used towards the calculation of composites but can be administered to better understand the child's motor functioning.

MOTOR FUNCTION

Motor Tests

SUBDOMAIN	MEASURE	AGE IN MONTHS	DESCRIPTION	MODALITY	ADMIN TIME IN MINUTES
Gross Motor	Get Up and Go	1-42	An assessment of postural transitions, locomotion, and overall gross motor ability by observing how children get up, travel 3 meters, and get onto and off a step platform.	Observation	4.5
Fine Motor	Reach to Eat	1-42	An assessment of fine motor dexterity, manual planning, and fine motor ability by observing how children grasp small objects and use tools.	Observation	5.5
Postural Control (Gross Motor, supplementary)	Sit and Stand	1-42	An assessment of stationary postural control by observing children who are placed in various sitting and standing positions.	Observation	3.5

Note: All NBT Tests require additional equipment and/or manipulatives to properly administer.

NIH Baby Toolbox®
www.nihbabytoolbox.org

Principal Investigator

Richard C. Gershon, PhD
Chief of Outcome and Measurement Science,
Department of Medical Social Sciences
Feinberg School of Medicine
Northwestern University

NIH Project Manager/COR

Kathy Mann Koepke, PhD
Child Development and Behavior Branch
National Institute of Child Health and Human
Development/National Institutes of Health

Sponsors

Primary support provided through the NIH Blueprint for Neuroscience Research. Blueprint is a collaborative framework that includes the NIH Office of the Director together with NIH Institutes and Centers that support research on the nervous system.

This project has been funded in whole or in part with Federal funds from the Blueprint for Neuroscience Research and the NICHD, National Institutes of Health, Department of Health and Human Services, under Contract No. 75N94019D00005.