Developmental Cognitive Neuroscience 70 (2024) 101461

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

journal homepage: www.elsevier.com/locate/dcn

Measurement of emerging neurocognitive and language skills in the HEALthy Brain and Child Development (HBCD) study

Julie A. Kable [a,*], Alexandra S. Potter [b], Natacha Akshoomoff [c], Patricia M. Blasco [d], Stefanie C. Bodison [e], Lucia Ciciolla [f], Sherry DeGray [g], Zoe Hulce [g], Emily S. Kuschner [h], Britley Learnard [g], Monica Luciana [i], Alexandra Perez [j], Miriam A. Novack [k], Tracy Riggins [l], So Yeon Shin [m], Sidney Smith [n], Jennifer Vannest [o], Eric.H. Zimak [p], the HBCD Neurocognitive and Language (NCL) Workgroup

[a] Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA 30329, United States
[b] Clinical Neuroscience Research Unit, Department of Psychiatry, 1 South Prospect Street Arnold 6, Burlington, VT 05401, United States
[c] Department of Pediatrics, UC San Diego, La Jolla, CA, United States
[d] Department of Pediatrics, School of Medicine, Institute on Development & Disability, Oregon Health & Science University, United States
[e] Department of Occupational Therapy, College of Public Health and Health Professions, University of Florida, Gainesville, FL 32611, United States
[f] Department of Psychology, Oklahoma State University, 116 Psychology Building, Stillwater, OK 74074, United States
[g] Department of Psychiatry, Clinical Neuroscience Research Unit, University of Vermont, Burlington, VT 05401, United States
[h] Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Scientist and Licensed Psychologist, Departments of Radiology and Psychiatry, The Children's Hospital of Philadelphia Philadelphia, PA 19146, United States
[i] Department of Psychology, University of Minnesota, Minneapolis, MN 55455, United States
[j] Department of Psychiatry & Behavioral Sciences Emory University School of Medicine, Atlanta, GA 30329, United States
[k] Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, United States
[l] Department of Psychology, 4094 Campus Drive, University of Maryland, College Park, MD 20742, United States
[m] Department of Human Development and Quantitative Psychology, University of Maryland, College Park, MD 20742, United States
[n] Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA 30329, United States
[o] Department of Communication Sciences and Disorders, University of Cincinnati, Speech-Language Pathologist, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States
[p] Department of Psychiatry and Behavioral Sciences, University of New Mexico Health Sciences Center, Albuquerque, NM 87131, United States

ARTICLE INFO

Keywords:
HBCD
Cognitive development
Language
Protocol
Executive function

ABSTRACT

The HEALthy Brain and Child Development (HBCD) study, a multi-site prospective longitudinal cohort study, will examine human brain, cognitive, behavioral, social, and emotional development beginning prenatally and planned through early childhood. The study plans enrolling over 7000 families across 27 sites. This manuscript presents the measures from the Neurocognition and Language Workgroup. Constructs were selected for their importance in normative development, evidence for altered trajectories associated with environmental influences, and predictive validity for child outcomes. Evaluation of measures considered psychometric properties, brevity, and developmental and cultural appropriateness. Both performance measures and caregiver report were used wherever possible. A balance of norm-referenced global measures of development (e.g., Bayley Scales of Infant Development-4) and more specific laboratory measures (e.g., deferred imitation) are included in the HBCD study battery. Domains of assessment include sensory processing, visual-spatial reasoning, expressive and receptive language, executive function, memory, numeracy, adaptive behavior, and neuromotor. Strategies for staff training and quality control procedures, as well as anticipated measures to be added as the cohort ages, are

* Correspondence to: Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, 12 Executive Park, Atlanta, GA 30329, United States.

E-mail addresses: jkabl01@emory.edu (J.A. Kable), alexandra.potter@uvm.edu (A.S. Potter), nakshoomoff@health.ucsd.edu (N. Akshoomoff), blascopa@ohsu.edu (P.M. Blasco), stefaniebodison@phhp.ufl.edu (S.C. Bodison), lucia.ciciolla@okstate.edu, lucia003@umn.edu (L. Ciciolla), sherry.degray@uvm.edu (S. DeGray), zoe.hulce@uvm.edu (Z. Hulce), kuschnere@chop.edu (E.S. Kuschner), britley.learnard@uvm.edu (B. Learnard), alexandra.perez2@emory.edu (A. Perez), miriam.novack@northwestern.edu (M.A. Novack), riggins@umd.edu (T. Riggins), syshin@umd.edu (S.Y. Shin), sidney.elise.smith@emory.edu (S. Smith), vannesjr@ucmail.uc.edu (J. Vannest), ezimak@salud.unm.edu (Eric.H. Zimak).

https://doi.org/10.1016/j.dcn.2024.101461
Received 7 March 2024; Received in revised form 12 September 2024; Accepted 25 September 2024
Available online 28 September 2024

reviewed. The HBCD study presents a unique opportunity to examine early brain and neurodevelopment in young children through a lens that accounts for prenatal exposures, health and socio-economic disparities.

1. Introduction

The HEALthy Brain and Child Development Study (HBCD) is a National Institute of Health (NIH) supported study enrolling over 7000 participating families from 27 sites across the United States and following them from pregnancy through childhood. The study's focus is to assess human brain development and factors that influence it, including potential teratogens and postnatal threats such as exposure to parental psychopathology, maltreatment, or other social adversities. The overall scope and aims of the project are outlined in the first paper of this special issue (Nelson et al., 2024). Data collection began in July 2023, and the first curated data release for this large-scale open-science project is anticipated in 2025. Information about the assessments used in this study is critical so researchers can prepare to utilize the information upon its first scheduled data release.

The overall organization of the HBCD Study Consortium is complex with a multi-faceted system of checks and balances to assure that the assessment scheme is optimized. The HBCD study protocols are proposed by assessment workgroups and are then reviewed by a Diversity, Equity and Inclusion Committee for feedback. Assessment protocol proposals are next presented to the Council of Investigators, and finally voted on by the Consortium's Steering Committee to result in the final protocol. This manuscript focuses on the activities of the Neurocognition and Language Workgroup (NCL). The workgroup was charged with proposing developmentally sensitive, longitudinal assessments of key language and neurocognition constructs that minimize participant burden, are psychometrically sound, are appropriate for the diverse families participating in the HBCD study, and are available in both Spanish and English languages.

Evaluations of neurocognitive and language development provide insights into functional outcomes of brain maturation. Neurocognitive domains of interest delineated by the workgroup were as follows: sensory processing, visual-spatial reasoning, expressive and receptive language, executive function, memory, numeracy, adaptive behavior, and neuromotor. This manuscript outlines the considerations that were taken in selecting the measures, the psychometric characteristics of the selected measures, and details regarding data collection methods. The manuscript additionally describes procedures for training study staff, including steps taken to maintain reliability over the course of the study.

2. Proposed research methods

2.1. Measurement selection considerations

In selecting measures, both caregiver report of the child's development and direct observation of skills using standardized assessment procedures were prioritized when possible to minimize biases. Several domains of sensorimotor function, neurocognition, language and memory will be obtained using the newly developed NIH Infant and Toddler (Baby) Toolbox (NBT) (NIH Infant and Toddler Baby Toolbox, 2022; Gershon et al., 2024). The NBT is a new developmental assessment designed for infants from 0 to 42 months and administered using an iPad tablet. The iPad technology offers innovative measurement techniques, including looking-time tasks made possible via automated eye-tracking for assessing infants and touchscreen options for assessing toddlers. The ease of training was also a consideration in adopting the NBT measures as the staff available to administer the protocol at the HBCD study sites varied in early childhood assessment experience, ranging from infant assessment specialists to research assistants with no previous standardized infant assessment experience. Select subtests from the NBT are administered at all in person study visits starting when participants are 3–9 months of age and continues throughout the early preschool period. The NBT was under development during protocol construction and early pilot testing. The efficiency of the NBT and the strong scientific evidence (NIH Infant and Toddler Baby Toolbox, 2022) to support the language and neurocognition tasks included in the NBT supported the workgroup's decision to rely heavily on the NBT for longitudinal assessment of language and cognition. However, there was no available information on the psychometric characteristics of the NBT subtests selected for inclusion in the HBCD study at this time. Despite this limitation, the workgroup decided to move forward with a beta version of the NBT that was made available for training research staff and collecting data until the public version is released. The workgroup plans to review the NBT data as they become available and make any necessary protocol adjustments.

Given the unknowns associated with the NBT, the workgroup opted to include a one-time assessment using the Bayley Scales of Infant and Toddler Development, 4th edition (Bayley-4) (Bayley and Aylward, 2019a) between 9 and 15 months to sample important domains identified by the workgroup. The Bayley-4 was selected as it is well-normed and validated (Bayley and Aylward, 2019b) and its previous versions are sensitive to the impact of exposures to substances and other environmental, social and biological factors on development (Amoros et al., 2019; Beker et al., 2023; Bakhireva et al., 2019; Lowe et al., 2019). Recent formulations of the Bayley-4 suggest that it may be able to characterize early variations in the development of executive functioning skills (Aylward et al., 2022). Given that the Bayley-4 is an established clinical measure, it allows the consortium an opportunity to provide feedback to families about their child's developmental status. Site clinicians will provide caregiver resources for community assessment and early intervention services as necessary or provide reassurance to families whose children did not have developmental concerns.

As the HBCD study is enrolling children from English and/or Spanish speaking homes, all assessments are offered in both languages. The infant's language exposure will be screened both to select the appropriate administration language and minimize challenges to the sites during the family's in-person assessment experience (DeAnda et al., 2016). Caregivers select their preference for the language used during their child's upcoming visit and sites are encouraged to provide multilingual assessors when possible for children with substantial multilingual exposure. The caregiver also completes a questionnaire to quantify their child's exposure to different languages. Data from this measure will be available with the public data release.

All caregiver questionnaires (The Sensory Processing Measure – 2nd Edition (SPM-2; (Parham et al., 2021)), Vineland Scales of Adaptive Behavior, 3rd Edition Parent/Caregiver Report (Vineland-3, (Sparrow et al., 2016)), and MacArthur-Bates Communicative Development Inventories (MB-CDIs*)* (Fenson et al., 2007, 1993; Marchman et al., 2023; Marchman and Dale, 2023) have both English and Spanish versions available. The NBT was also developed in both English and Spanish. The Bayley-4 was not available in Spanish despite a recent translation being available for the Bayley Scales of Infant and Toddler Development, 3rd edition (Bayley, 2023). The language translation from the Bayley, 2023 Spanish translation was used to guide the translation of the HBCD study Bayley-4 as a result of the substantial item overlap. The translation was initiated by HBCD investigators and staff who had extensive prior testing experience with the Bayley and were Spanish speaking. The translated documents were then reviewed by independent reviewers certified in Spanish language translation who completed the back translation. Additional edits were made by the first team and reviewed again by the second team to reach consensus. A document with the Spanish translation of instructions and caregiver queries was created for administration. Materials were constructed so that the Spanish stimulus book instructions could be inserted into the primary stimulus book directly under the English instructions. Spanish versions of the Bayley-4 child book (*Clifford's Animal Sounds*) used during testing were required to be purchased. The Deferred Imitation task was translated using a similar review process.

2.2. *Timeline and protocol*

Data collection timepoints are outlined in Fig. 1 up to V05. This manuscript discusses these timepoints and additional NCL assessments planned to capture important development changes in the proposed functional domains. There are five in person visits (Visit 1 (V1), Visit 2 (V2), Visit 3 (V3), Visit 4 (V4), Visit 6 (V6)) and two remote visits (Visit 5 (V5) and Visit 7 (V7), each occurring within 1 month of the previous in-person visit. Relative to infant age, V1 is conducted prenatally, V2 is conducted from 0 to 1 month of age, V3 is from 3 to 9 months of age, V4 is from 9 to 15 months of age, and V6 is from 15 to 30 months of age. Table 1 details the protocol for NCL Workgroup including domains of interest, measures used, and assessment timepoints. Fig. 2 identifies the methods of assessment, parent report or direct observation, by each domain across the timepoints. Assessments of neurocognition and language are initiated in V3 when the infant is between 3 and 9 months of adjusted age and are incorporated throughout the remaining visits. At the time of this writing, the protocol is being piloted so modifications may occur as result of the outcomes of the pilot. Data collection methods for each of the domains of assessment are delineated below.

2.2.1. *Sensory processing*

Sensory processing refers to the brain's ability to integrate auditory, gustatory, olfactory, visual, tactile, vestibular, and proprioceptive information from one's own body and the environment (Dunn et al., 2022). At birth, the infant's nervous system is bombarded by sensory data that must be processed and integrated to make sense of the world. Sensory processing serves as the foundation on which higher order skills are built. In children, the development of self-regulation, social relatedness, and motor abilities specifically relies on efficient and effective processing of sensory information as the child interacts with the environment and people around them. Deficits in sensory integration can cause delays in neurodevelopment and lead to functional deficits in play skills, fine and gross motor development, praxis, social interaction, and

Table 1
Neurocognition and Language Domains and Measures in the HBCD Study.

Construct	Measure	Assessment Timepoints
Sensory processing	Infant and Toddler Sensory Processing Measure, 2nd edition (Parham et al., 2021)	V3[a], V5
Visual-Spatial Reasoning	Bayley−4[b] Cognitive Cluster	V4
	NBT[c] Visual Reception	V6
Language	Bayley−4 Language Cluster (Expressive & Receptive Communication)	V4
	MacArthur-Bates Communicative Development Inventories (MB-CDIs) (Fenson et al., 2007, 1993; Marchman et al., 2023; Marchman and Dale, 2023)	V4 and every six months until 30 months of age.
	NBT Looking-While-Listening	V4, V6
	NBT Mullen Receptive and Expressive Language	V3,V6
	NBT Picture Vocabulary	V6 (>25 months)
Executive Function	EF Factor Score from Bayley−4 (Aylward et al., 2022)	V4
	NBT Executive Function	V4, V6
Memory	NBT Executive Function: Visual Expectation	V4, V6
	NBT Executive Function: Visual Delayed Response	V4, V6
	Deferred Imitation Task	V4, V6
	NBT Toolbox Picture Sequence Memory	V6 (>25 months)
Numeracy	NBT Numerical Change Detection	V4, V6
	NBT Object Counting	V6 (>25 months)
	NBT Subitizing	V6 (>25 months)
	NBT Verbal Arithmetic	V6 (>25 months)
	NBT Who has more?	V6 (>25 months)
Adaptive Function	Vineland Adaptive Behavior Scale, 3rd edition (Sparrow et al., 2016)	30–42 months
Neuromotor Function	Bayley−4 Motor Cluster	V4
	Vineland Adaptive Behavior Scale, 3rd edition-Motor Domain	30–42 months

[a] V3 refers to Visit 3, V4 refers to Visit 4, V5 refers to Visit 5, V6 refers to Visit 6, and V7 refers to visit 7

[b] Bayley 4: Bayley Scales of Infant and Toddler Development, 4th edition (Bayley-4) (Bayley and Aylward, 2019a)

[c] NBT: NIH Infant and Toddler (Baby) Toolbox (NBT) (NIH Infant and Toddler Baby Toolbox, 2022; Gershon et al., 2024).

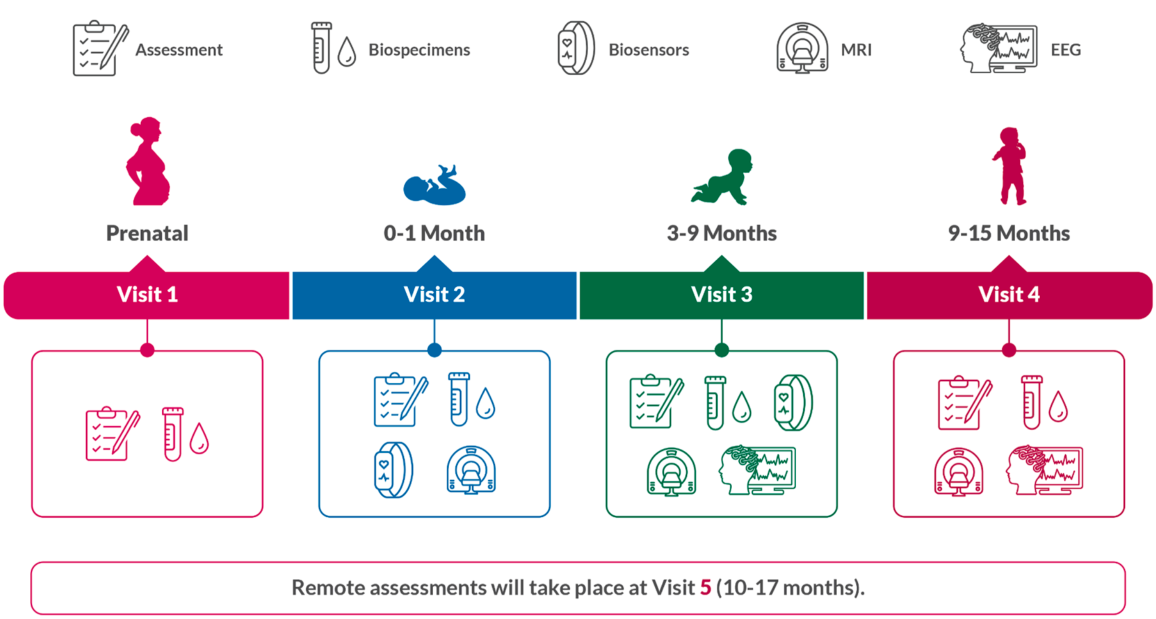

Fig. 1. Timeline of the HBCD Study's Data Collection.

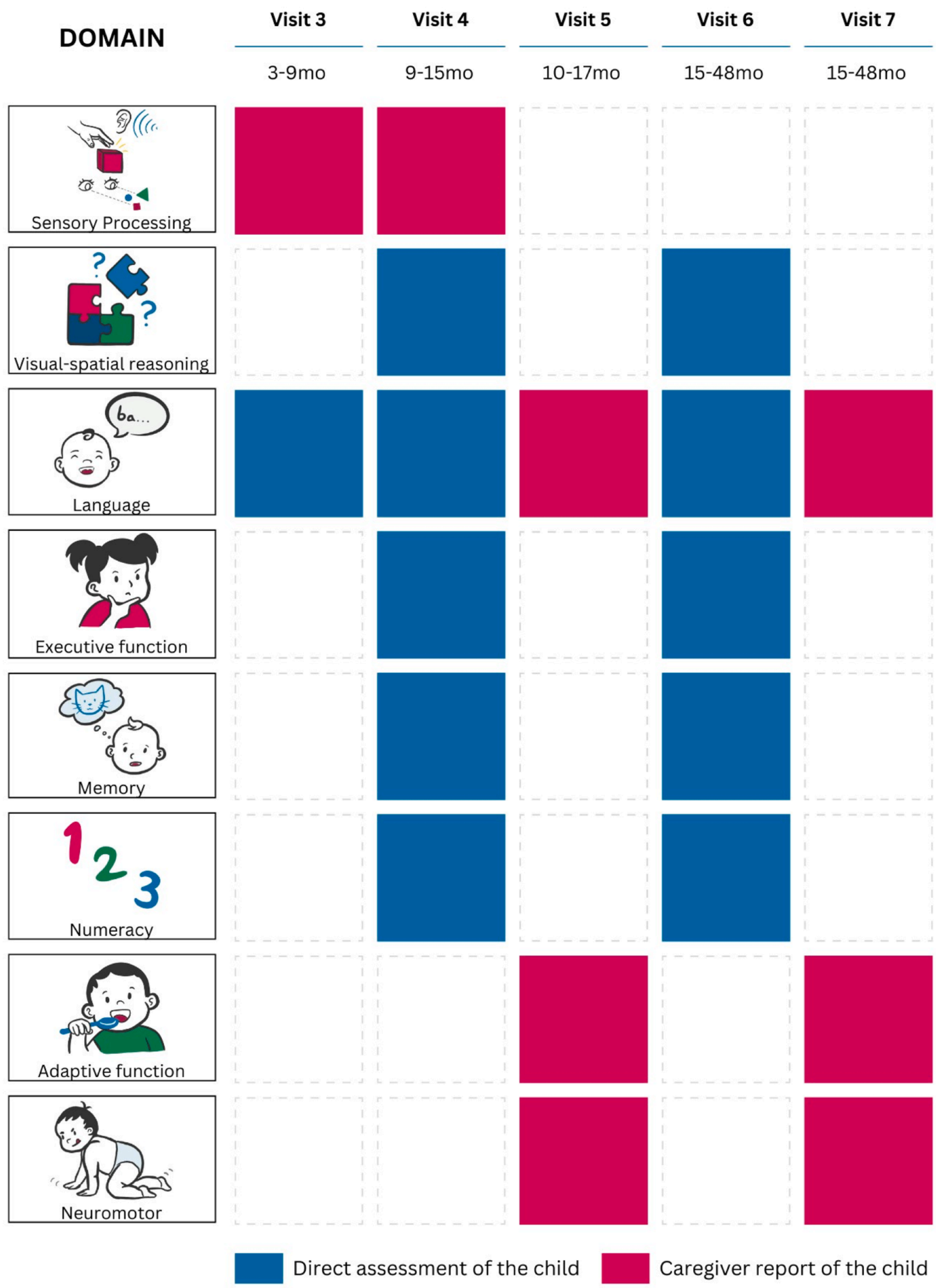

Fig. 2. Domains of Assessment by Method of Data Collection and Visit (V) Schedule.

self-regulation (Dunn, 2007). Several neurodevelopmental conditions present with co-occurring sensory processing deficits, including autism spectrum disorders (Patil and Kaple, 2023; Baum et al., 2015; Narzisi et al., 2022), attention deficit disorders (Kern et al., 2015), and fetal alcohol spectrum disorders (Fjeldsted and Xue, 2019).

In the HBCD study, the sensory processing abilities of the child will be assessed at V3 and V5 using the Infant/Toddler age level forms of the SPM-2. The SPM-2 is a standardized, norm-referenced questionnaire that assesses the sensory domains of vision, hearing, touch, taste, smell, body awareness, balance and motion, planning and ideas, and social participation in individuals ages 4-months to 87-years using a 4-point Likert scale. The caregiver completes the Infant/Toddler/Preschool age level forms of the SPM-2. They were standardized on 1303 typically developing 4-month to 5-year-old children and have internal consistency (α) estimates ranging from .79 to .98 and test-retest reliability (r) estimates ranging from .82 to .96. The SPM-2 questionnaires may be completed in person or remotely and is commercially available in both English and Spanish forms. The SPM-2 has been used with several different neurodiverse groups (Parham et al., 2021; Narzisi et al., 2022; Jirikowic et al., 2013) and diverse cultural groups (Román-Oyola and Reynolds, 2013; Fernández-Pires et al., 2021; Amer et al., 2022).

2.2.2. *Visual-spatial reasoning*

Visual-spatial reasoning is a complex set of neurocognitive skills that impact a range of developmental skills (Johnson and Moore, 2020; Davis, 2015; Blazhenkova and Kozhevnikov, 2010). In infants and toddlers, this domain of functioning is often assessed via gaze tasks which encode variations in visual attention to stimuli or through sensorimotor tasks that involve the manipulation of objects. The latter allows for the assessment of early sensorimotor coordination needed to explore objects, visual perception skills, object permanence, problem-solving, magnitude estimation and comparisons, sequencing, color recognition, visual and spatial memory and aspects of imaginative play. The HBCD study battery includes the NBT Mullen Visual Reception, a sensorimotor-based subtest derived from items selected from the Mullen Scales of Early Learning (MSEL; Visual Reception). The original Visual Reception test of the MSEL has been used frequently in developmental research and has been found to be predictive of later cognitive status and developmental gains related to intervention, particularly in children with autism spectrum disorders (Kauley et al., 2024; Swineford et al., 2015; Bishop et al., 2011). The NBT version of this test involves a subset of the original items (n = 28) adaptively administered in developmental order of item difficulty, and automatically determined by the NBT application, reducing administrator burden. For older toddlers and preschoolers, the NBT Visual Reception task includes a visual matching component. This involves using the touchscreen features of the iPad to select from an array of four pictures of which one matches a picture presented at the top of the screen.

At V4 (9–15 months), the Bayley-4 Cognitive Cluster (Bayley and Aylward, 2019a) will also be used to assess sensorimotor-based visual-spatial learning. The Cognitive Cluster is reported as both a subtest score with a mean of 10 and standard deviation of 3 or as a standard score with a mean of 100 and a standard deviation of 15 points. From the technical manual of the Bayley-4 (Bayley and Aylward, 2019b), test-retest reliability of the Cognitive Cluster ranged from .80–.83 (Corrected r values) and internal consistency ranged from .90 to .98 with individuals 0 to 42 months of age. Correlation with the Cognitive Cluster of the Bayley-III was .70 (corrected R) and the Wechsler Preschool and Primary Scale of Intelligence, IV (WPPSI-IV) was .79. Scores on this scale have been found to differentiate children with an autism spectrum disorder, developmental delay, prematurity and a history of prenatal alcohol and drug exposure from matched controls (Bayley and Aylward, 2019b). The various forms of the Bayley have been used in numerous developmental studies (Dack et al., 2022; Biagas et al., 2023; Del Rosario et al., 2021; Subramoney et al., 2018) across diverse populations and has been identified as the *gold standard* for infant assessment (Oyungu et al., 2022; Marban-Castro et al., 2022; Coles et al., 2019).

2.2.3. *Language*

Trajectories of language development in infancy and early childhood are well-described as is the importance of early language acquisition for later developmental outcomes. Language measures in the HBCD study include both normed, well-established measures (the MB-CDIs and language sections of the Bayley-4) and more novel measures [NBT tasks of Looking While Listening, the NBT Mullen Expressive (Observational and Prompted) and the NBT Mullen Receptive Language subtests adapted from the original MSEL language tasks (Mullen, 1991)], as well an adaptation of the National Institute of Health's Toolbox (NIHTB) Picture Vocabulary Test (TPVT) (Gershon et al., 2014), which has been tested with children 3 and older (Gershon et al., 2013). The NBT is extending the normative information on this subtest down to 2 years.

The MB-CDIs is first administered at V5 and will be administered 3 times between ages 10 – 30 months to assess developmental change in language skills during this critical phase of language development. Specifically, caregivers are presented with a list of words, phrases, and gestures and then asked to indicate whether their child uses and/or understands these specific items. The Words and Gestures (CDI:WG) form (8–18 months) is focused on early words, phrases and gestures. The Words and Sentences (CDI:WS) form (16–30 months) also assesses children's use of word combinations and grammatical markers. In the HBCD study, infants who are less than 16 months of age will do the CDI: WG and infants greater than or equal to 16 month of age will be given the CDI:WS. Summary scores, percentiles, and individual item responses are generated from the CDIs. In addition, norms for the American English version of the CDIs have recently been updated to include a larger normative sample better reflecting U.S. demographics (Marchman and Dale, 2023). For Spanish-speaking families in the HBCD study, the Mexican Spanish version of the CDIs (Jackson-Maldonado et al., 2003) will be administered (see below for specifics on language of administration). The CDIs have been shown to be sensitive to changes in language development associated with a number of factors affecting the first year of life, including premature birth (Foster-Cohen et al., 2007; Cattani et al., 2010; Charkaluk et al., 2019) and prenatal exposures (Olesen et al., 2018; Hernandez-Martinez et al., 2017; Doherty et al., 2019). Because the CDI involves reporting on the child's production and comprehension of specific items in their native language(s), the decision for which CDI to administer will be based on the language(s) that the child is most exposed to, as the language(s) will be most reflected in their early vocabulary. As described above, caregivers will be asked to report their child's language exposure. Families whose children are exposed at least 30 % of the time to English or Spanish (or both) will be asked to complete the CDI in that language (or both).

In most cases, the reporting caregiver is likely to speak the same language to which the child is exposed. For bilingual or multi-lingual children, different caregivers may have various levels of proficiency in each of the languages that the child is learning and may communicate with the child in one or more languages most frequently. Given the scope of the HBCD study, it is not feasible to require responses from multiple caregivers. Therefore, with each administration of the CDI, we will document who completes the report, which language they speak most frequently with the child, and a rating of how confident they feel reporting on the child's use of that language. This information will allow researchers who access the HBCD study dataset to include or exclude cases in analyses where the reporter was not confident about responses in a particular language.

The NBT Mullen Receptive and Expressive Language subtests will be assessed during in person visits conducted at V3 and V6. These subtests were derived from a subset of items from the MSEL and were obtained by observing the child's spontaneous vocalizations and responses to vocal/verbal prompts and challenges. The NBT Looking While Listening, which is based on the widely used visual world paradigm (Allopenna et al., 1998), relies on infants' tendency to direct their gaze toward a visual referent (e.g., a car) when hearing that referent's label in natural speech (e.g., *Look at the car!)*. Item difficulty levels were matched for the Spanish version rather than simple literal translations of the English language words used in the task (personal communication with NBT development staff). For infants 25-months and older, the NBT includes a Picture Vocabulary Task where the child is shown four pictures and asked to point to the item verbalized.

The Bayley-4, which is only administered at the 9–15 months' time window (V4), will provide a standardized assessment of receptive and expressive communication skills. Each of the subtests are reported as a scaled score with a mean of 10 and a standard deviation of 3. The subtests are then aggregated to formulate a Language Cluster with a mean of 100 and a standard deviation of 15 points. From its technical manual, corrected correlations of test-retest reliability ranged from .80–.87 and internal consistency estimates ranged from .81 to.99 across 0–42 months. The language components of the Bayley were found to be predictive of scores on later preschool language tests (Preschool Language Scale 4th edition (Zimmerman et al., 2002)) and WPPSI-IV Full Scale IQ (Bayley and Aylward, 2019b) and have been found to differentiate children with neurodevelopmental conditions and known history of prenatal alcohol and drug exposure from matched controls (Bayley and Aylward, 2019b). Earlier versions of the Bayley have been used in numerous developmental studies (e.g., (Dack et al., 2022); Biagas et al., 2023; Del Rosario et al., 2021; Subramoney et al., 2018).

2.2.4. *Executive functioning*

Executive functions (EFs) are related neural processes that make it possible for us to use focused attention, reason and problem solve; make choices, engage in self and emotional control (inhibit), use working memory, understand multiple perspective, use cognitive flexibility to shift between attributes, and plan/organize our daily routines (Diamond, 2020). The prefrontal cortex (Kfir et al., 2009; O'Hare et al., 2009; Fryer et al., 2007; Warren et al., 2011; Olegard et al., 1979; Sowell et al., 2007) and the connectivity of the prefrontal cortex to other brain regions (Wozniak et al., 2013) has been found to be the neural substrate from which EF skills arise. Changes in brain maturation (Bell, 1998), particularly myelination of the prefrontal cortical areas of the brain, over the course of the first two years of life may contribute to differences in EF skills. These brain maturation changes correspond to developmental maturational processes allowing for effortful control over the regulation of attention and capacity to sustain mental effort (Ruff and Rothbart, 1996) needed to initiate goal attainment behaviors. Meta-analytic or systematic review findings indicate that prenatal alcohol exposure (Khoury et al., 2015), childhood maltreatment (Lund et al., 2020), and maternal depression (Lund et al., 2020) are among the prenatal and postnatal variables associated with poorer executive function in children.

Both the NBT and the Bayley-4 will contribute to the assessment of this domain. The NBT Executive Function task is a 10-minute measure that uses eye gaze tasks to assess early learning and EF skills. The Bayley-4 includes many developmental tasks which many tap into aspects of EF. The EF components of the Bayley-4 were identified in a previous investigation (Aylward et al., 2022) using factor analysis of these items and these will also be available to end users of the HBCD study dataset.

2.2.5. Memory

Memory is a cornerstone ability upon which children build knowledge and understanding of the world. For example, memory guides children's attention, it is required for learning new words, concepts, skills, and it is essential for distinguishing familiar faces and building social relationships. Memory impairments can interfere profoundly with a child's life, particularly in relation to their schooling, daily routines, life satisfaction and success. Due to the prolonged development of memory and its reliance on neural circuitry susceptible to environmental influences (e.g., the hippocampus with high density of glucocorticoid receptors), memory shows great plasticity (Nelson, 2007). Environmental factors both promote the development of memory (e.g., narrative style, strategy use) as well as hinder it (e.g., stress, nutrient deficiencies, prenatal exposure to substances). Assessing memory in the HBCD study is essential, as it will shed light on how this cornerstone ability is used by infants to build up knowledge of the world; how it interacts with other neurocognitive, social and emotional developmental abilities; and how exposure to a variety of circumstances results in individual differences.

The HBCD study will assess both short- and long-term memory using the subtests from the NBT Executive Functioning task and the Deferred Imitation Task (DI). The NBT Executive Functioning gaze measure includes portions that assess recognition memory (memory after a 6 minute delay for objects previously learned during the habituation portion). The same tasks are adapted into touchscreen versions for toddlers. The Memory Task Learning test is a touch-based assessment of immediate memory, where children are asked to recognize stimuli presented on the trial immediately before. The Memory Task Test is a touch-based assessment of delayed memory, in which children are tested on their recognition of individual items from the Memory Task learning portion after a 6–8 minute delay. The NTB Executive Functioning battery also includes two versions of a Visual Delayed Response Task. In the gaze-version for infants, an object appears and is then occluded behind one of two boxes displayed on each side of the screen. A curtain occludes the screen for a brief delay, after which the boxes again become visible and child's looking to the old location is assessed. A similar touch-based version is administered to older toddlers. For older toddlers and preschoolers, the NTB Picture Sequence Memory Form Task (Dikmen et al., 2014) will also be administered. This task assesses episodic memory by presenting a series of pictures in order with an audio description. The pictures are scrambled, and participants are asked to arrange the pictures in the order in which they were presented.

Deferred imitation (DI) is a "tried and true" robust paradigm that dates back to Jean Piaget (Piaget, Jean (ed), 1962) and involves an infant or young child observing a task and then imitating the task at another time. Performance on the DI task has been shown to be sensitive to alterations in memory in infants due to multiple risk-factors, including: prenatal iron deficiency, preterm birth, history of institutional rearing, and neglect (see (Bauer, 2010) for review) and predicts memory performance at school age (Riggins et al., 2013). Finally, DI avoids issues with the ambiguity of using looking time measures alone to assess memory, especially across development (e.g., (Richmond et al., 2007)). DI will be assessed at V6. DI assesses long-term item memory and ordered recall. Participants are shown a series of novel actions and, after a delay, encouraged to imitate them. Two different novel event sequences, created using custom 3D printed materials, will be shown (Dr. Cheatham, UNC). Participants a will receive two tasks with one including 3-step sequences and the second including 4-steps. This assessment yields a measure of long-term memory (> 10 min) for individual actions (i.e., item memory) and for actions in the correct temporal order (i.e., ordered recall). DI taps the same type of memory assessed by the NIH ToolBox in early childhood (i.e., Picture Sequencing Task (Weintraub et al., 2013a)) but is able to be administered earlier in development.

2.2.6. Numeracy

Although there are components of early numeracy on the NBT and Bayley-4, they are embedded within the assessment of visual-spatial reasoning. There is clear overlap between visual-spatial reasoning and later math ability (Rittle-Johnson et al., 2019; Lauer and Lourenco, 2016). In this context, we will only discuss those tasks specifically designed to assess numerical perception and numeracy. The Numerical Change Detection task is the earliest task in the NBT that assesses infants' perception of numeracy and is administered to infants ages 6–24 months. The task utilizes gaze tracking in a change detection paradigm to assess the approximate number system. Participants are presented with two concurrent streams, one of which alternates between numerical quantities and one of which shows a constant quantity. In the preschool period, several number-related tasks are included in the NBT. Object and Verbal Counting assesses counting and cardinality by asking children to count as high as possible and to report the number of objects presented on a screen. Subitizing assesses approximate number system and number recognition by asking children to report on the number of objects presented briefly on a screen. Verbal Arithmetic assesses addition and subtraction by asking children to respond to simple arithmetic questions verbally. Finally, Who Has More, a touch-screen paradigm, assesses number approximation by asking children to select which of two numerical quantities on a screen is larger.

2.2.7. Adaptive function

Adaptive function is measured in the HBCD study as it assesses how an individual is able to function in the real world (Sparrow et al., 2016). The Vineland-3 is a caregiver report of functioning in the domains of communication, daily living, socialization, and motor functioning. Whereas ability measures focus on what someone can do in a testing situation, the Vineland-3 focuses on what they do in daily life. The Vineland-3 is a revision of the widely used adaptive functioning measure (Sparrow et al., 2005). The HBCD study uses the Parent/Caregiver Form, which is administered electronically. Caregivers are sent a link that allows them to respond to items via Pearson's electronic scoring system and the results are transmitted to the data administrative core of the HBCD study. The Vineland-3 revision (Sparrow et al., 2016) is well-normed with internal consistency $\geq$.90 in all areas of the Parent/Caregiver Form. Test-retest reliability scores were showed corrected r's for infants between 0 and 2 years of age ranging from.71 to.90 for summary scores and from .74 to.91 for subdomain scores. For children ages 3–6, corrected r's ranged from .83 to.91 for summary scores and .60

to.90 for subdomains. Vineland-3 scores obtained are well-correlated with those obtained from the previous version of the task and continue to differentiate clinical groups (i.e., intellectual disability, autism spectrum disorders) (Sparrow et al., 2016). This measure is anticipated to be initially administered when when infants are 30–42 months. The test has subdomains that are reported as v-scale scores with a mean of 15 and a standard deviation of 3. Subdomain scores are then aggregated into the four skill areas to generate standard scores with a mean of 100 and a standard deviation of 15 points (Communication, Daily Living Skills, Socialization, and Motor Skills) and an overall composite score (Adaptive Behavior Composite).

2.2.8. Neuromotor function

Disruption to early motor functioning is often an important indicator of alterations to brain development (Doherty et al., 2019; Hendricks et al., 2020). Assessments of neuromotor functioning are typically subdivided into fine motor and gross motor skill assessments. Fine motor skills incorporate assessments of the child's early grasping skills, their visual-perceptual integration of grasping and reaching, their motor planning skills, and the speed with which they are able to carry out these functions. Gross motor skills measure the child's head and trunk control and movement of their limbs and upper body. Impaired neuromotor functioning has been linked to prenatal exposure to lead (Despres et al., 2005), alcohol (Hendricks et al., 2020; Lucas et al., 2014), and opioids (Mahabee-Gittens et al., 2024; Yeoh et al., 2019), suggesting this domain of functioning was important to capture in the HBCD study.

In the HBCD study, the Bayley-4 Motor Cluster will be administered at V4. The assessment includes a summary score, with a mean of 100 and standard deviation of 15, and subscale scores of fine and gross motor development, with a mean of 10 and standard deviation of 3. From its technical manual, corrected correlations of test-retest reliability ranged from .71–.88 and internal consistency estimates ranged from .84 to.97 across 0–42 months. The Vineland-3, given between 30 and 42 months, also has a measure of adaptive motor skills that includes the summary score with a mean of 100 and a standard deviation of 15 and subscales assessing adaptive fine and gross motor functioning. The combination of both of these measures will allow for integration of parental report and direct observation of neuromotor function that can be related to neuroimaging data collected in the study.

2.3. The HBCD study staff training and data collection quality control

2.3.1. Establishment of an infant testing core

An Infant Testing Core (ITC) was established to (1) ensure proficiency in administering the neurocognition and language measures; (2) support the training of staff and oversee processes for staff certification; and (3) maintain quality data collection over the course of the study. The core consists of investigators from five HBCD study sites who provide trainings and review video submissions of assessments to provide feedback and certification of the HBCD study staff who administer assessments. Each of these investigators has over 10 years of infant assessment experience. In addition, research assistants who were certified by either the NBT training team or by a core ITC investigator also review video submissions and provide feedback under the supervision of the ITC investigators. Forms were developed to provide feedback for staff to identify areas in need of improvement and a weekly meeting is held for the ITC to maintain consistency in the evaluation process.

Trainings are provided to the consortia research staff by experts in each assessment. Videos of the trainings and model assessments are available on a secure HBCD study platform for training. Manuals and standard operating procedures are maintained on the shared platform. Office hours for neurocognition and language assessments are provided weekly for the HBCD study staff to discuss issues with ITC investigators. Individual meetings are offered to provide individual feedback on certification videos.

The criteria for certification on each assessment includes three main components. The first is having and maintaining an adequate testing environment, including an environment free from distractions and availability of the appropriate physical furniture (i.e., tables, chairs, high-chair, and steps) necessary for the protocol. Second, assessors were required to demonstrate familiarity with the administration instructions, the assessment materials, and the iPad interface used by both the NBT and the Bayley 4. The third criteria focused on accuracy of data collection, including item administration and scoring; and adherence to procedures needed for gaze capture for the NBT gaze measures. Many traditional areas of concern that impact the quality of the assessments are minimized by the use of automated scoring software for both the NBT and Bayley-4. QR codes encode the child's study identification number, calculate the child's appropriate age and automate scoring. The automated scoring provides examiner feedback on attainment of basal and ceiling requirements during the administration and then performs calculations needed to obtain standardized summary scores, which are then electronically submitted to the Consortium's database.

2.3.2. The HBCD study consortium-wide staff training process

NBT: Research staff submit video recordings of their assessments, as well as item level scoring, and a self-evaluation form. Certification in NBT requires two separate video submissions. First, research staff submit video in which the staff member administers all of the non-gaze NBT measures as a role-play to an adult volunteer. This video is reviewed for appropriate environment, materials, and item-level administration and scoring. Feedback is provided to the assessor and site Principal Investigator. If an assessor does not pass they are required to repeat this step (administration to an adult).When the assessor passes, they move onto the next step of certification.

The second step of certification is a submission of NBT administration to an infant between 9 and 15 months of age, accompanied by the scoring and self-evaluation forms. Certifiers review the videos for administration and scoring errors, and provide comments and suggestions for improving efficiency and infant management. A failed certification requires resubmission of an infant NBT assessment. Once an assessor passes both rounds of certification, they are fully certified to administer the NBT.

Bayley-4: The HBCD study staff are required to complete a 12-hour on-line Pearson training and certification before evaluating an infant. Then they begin practice administrations and finally submit a recording of a Bayley-4 administration with an infant who is between 9 and 15 months of age, along with scoring and a self-evaluation form. These videos are reviewed by the ITC staff and, similar to NBT certification procedures above, item-level feedback on administration and scoring is provided. For the HBCD study, all Bayley-4 assessments are administered using QGlobal™, which is Pearson's online scoring platform. This platform facilitates administration for assessors as instructions and caregiver queries are readily available during the assessment process and the software tracks the basal and ceiling criteria needed to accurately assess the infant's performance. Similar certification criteria were used for Bayley-4 certification as described for the NBT certification.

Deferred Imitation (DI): A similar review process is used for DI, including submission of videos, self-assessments. Item level feedback is then provided for administration and scoring.

2.4. Monitoring data collection integrity over time

Once an assessor is certified, they can administer the NCL measures at study visits, and all assessors are required to participate in a peer reliability checking process. The peer-review process will be maintained for the duration of the HBCD study and will provide reportable reliability metrics on items and summary scores. These metrics will be continuously reviewed by the ITC and item retraining and/or site retraining will be driven by these metrics. Peer assessors will video record and upload their 10th administration of an assessment and every 20th administration after that. For each assessment they upload, they will complete a reliability check for a peer's video. This will involve reviewing the peer's video and completing item level scoring from the video. Peers can flag administration errors and/or questions for ITC review. The ITC will administer this process including scheduling and assigning videos for peer review and ensuring timely completion of the reviews.

3. Results

3.1. Data encoding and processing

Except for the language exposure assessment and the DI task, the neurocognitive and language measures used in the HBCD study involve entering the responses into third party software that is transmitted to the HBCD study Data Coordinating Center. The HBCD study administrative staff worked closely with the test publishers or developers to optimize the administration and scoring of the measures. As with all measures in the HBCD study, protected health information (PHI) is securely stored and managed within the HBCD study Data Coordinating Center. ID numbers and age are transmitted to the third party software using a QR code to protect privacy and to minimize errors. For measures completed by the caregiver, an electronic link that automatically provides information about the child's age at the time of completion will be generated. The responses will be securely transferred to the publishers web-based administration, scoring, and reporting systems and scores will be returned and uploaded to the consortium database. For the DI and language assessment, responses are entered directly into Longitudinal Online Research and Imaging System (LORIS), the data capture system used in the HBCD Study. All raw data and standardized scores are uploaded to the HBCD study secure, encrypted database and released as part of the annual curated data releases.

3.2. Convergence of the HBCD study with ABCD cognitive and language domains

The launch of the HBCD study is timely given its intersections with another NIH-funded consortium study, the Adolescent Brain Cognitive Development (ABCD©) study. ABCD is following a cohort of 11,878 adolescents, enrolled at ages 9–10, into young adulthood using a comprehensive developmentally-informed neuroimaging and behavioral battery that is sensitive to variations in experience, including substance use (Volkow et al., 2018). As of this writing, the ABCD participants are 15–16 years of age, and over 90 % of the sample has been longitudinally retained through seven years of assessment. ABCD assesses cognition at two-year intervals using a core set of measures administered using iPads, including the NIH Toolbox Cognitive Battery (Weintraub et al., 2013a), the Rey Auditory Verbal Learning Test (Rosenberg et al., 1984), and the Little Man Test of visuospatial cognition ((Luciana et al., 2018; Thompson et al., 2019), see Table 2 below). Other measures have been utilized intermittently to assess general intelligence (Matrix Reasoning Task (Wechsler, 2003)) as well as behaviors particularly relevant to adolescent decision-making such as risk-taking (Game of Dice Task (Brand et al., 2002)), peer influence over risk evaluations (Social Influence Task (Chawla et al., 2009)), and cognitive persistence under conditions of frustration tolerance (Behavioral Indicator of Resiliency to Distress (BIRD) Task (Daughters et al., 2009)). During non-scan years (i.e., no MRI), measures of delay discounting (see (Owens et al., 2022)), the emotional face-emotional word Stroop Task (Smolker et al., 2022), and the Stanford Mental Arithmetic Response Time Evaluation (Smarte; (Guillaume et al., 2023)) are administered. Despite obvious differences in developmental stage and response capacities for infants and preschoolers versus adolescents, the overlap in assessment constructs across the two studies is apparent.

The assessment of memory function across both studies serves as an illustrative example, where infant memory will be examined nonverbally through the deferred imitation paradigm, which is the same

Table 2
Alignment of Neurocognitive Constructs Measured by the HBCD and ABCD Studies.

Domain of Cognition	HBCD	ABCD
Sensory Processing	Infant and Toddler Sensory Processing Measure, 2nd edition (Parham et al., 2021)	N/A
Visual-Spatial Reasoning	Bayley−4[a] Cognitive Cluster NBT[b] Visual Reception	Little Man Task (Luciana et al., 2018) Wechsler Matrix Reasoning Test (Wechsler, 2003)
Language	Bayley−4 Language Cluster (Expressive & Receptive Communication) MacArthur-Bates Communicative Development Inventories (MB-CDIs) (Fenson et al., 2007, 1993; Marchman et al., 2023; Marchman and Dale, 2023) NBT Looking-While-Listening NBT Mullen Receptive and Expressive Language NBT Picture Vocabulary	NIH Toolbox Picture Vocabulary & Oral Reading (Gershon et al., 2014)
Executive Function	EF Factor Score from Bayley−4 (Aylward et al., 2022). NBT Executive Function	NIH Toolbox Flanker Task (Weintraub et al., 2013b) NIH Toolbox Pattern-Comparison Processing Speed Task (Weintraub et al., 2013b) NIH Toolbox Dimensional Change Card Sort Task (Weintraub et al., 2013b) NIH Toolbox List-Sort Working Memory Task (Weintraub et al., 2013b)
Memory	NBT Executive Function: Visual Expectation NBT Executive Function: Visual Delayed Response Deferred Imitation Task NBT Toolbox Picture Sequence Memory	NIH Toolbox Picture Sequence Memory (Weintraub et al., 2013b) Rey Auditory Verbal Learning Test (Rosenberg et al., 1984)
Numeracy	NBT Numerical Change Detection NBT Object Counting NBT Subitizing NBT Verbal Arithmetic NBT Who has more?	Stanford Mental Arithmetic Response Time (Guillaume et al., 2023)
Adaptive Function	Vineland Adaptive Behavior Scale, 3rd edition (Sparrow et al., 2016)	N/A
Neuromotor Function	Bayley−4 Motor Cluster Vineland Adaptive Behavior Scale, 3rd edition-Motor Domain	
General Intelligence	Bayley−4 Cognitive Bayley−4: Language Cluster NBT Visual Perception NBT Expressive and Receptive Language	NIH Toolbox Picture Vocabulary (Weintraub et al., 2013b) NIH Toolbox Oral Reading (Weintraub et al., 2013b) Wechsler Matrix Reasoning Test (Wechsler, 2003)
Peer Influence on Decision-Making	N/A	Social Influence Task (Chawla et al., 2009)
Cognitive Impulsivity/ Decision-making	N/A	Delay Discounting Task (Owens et al., 2022)
Emotional Influences on Conflict Monitoring	N/A	Emotional Face/Emotional Word Stroop Task (Smolker et al., 2022)
Risk-Taking Behavior	N/A	Game of Dice Task (Brand et al., 2002)

(continued on next page)

Table 2 (*continued*)

Domain of Cognition	HBCD	ABCD
Cognition Under Conditions of Frustration Tolerance	N/A	Behavioral Indicator of Resiliency to Distress (BIRD) Task (Daughters et al., 2009)

[a] Bayley 4: Bayley Scales of Infant and Toddler Development, 4th edition (Bayley-4) (Bayley and Aylward, 2019a)

[b] NBT: NIH Infant and Toddler (Baby) Toolbox (NBT) (NIH Infant and Toddler Baby Toolbox, 2022; Gershon et al., 2024).

hippocampally-based explicit memory process measured by the NIH Toolbox Picture Sequence Memory Task (Weintraub et al., 2013a). NIH Toolbox-based Picture Sequence Memory will be assessed in the HBCD study sample starting in the preschool period, and this task, combined with the Rey Auditory Verbal Learning Task, is foundational to the assessment of memory in ABCD. Using ABCD data, it has been shown that these memory processes improve with age, are hippocampally-mediated, are influenced by variations in youth socio-economic status, and are impacted by prenatal alcohol exposure (Lees et al., 2020; Anokhin et al., 2022; Botdorf et al., 2022). More nuanced integrations of findings across the HBCD study and ABCD will enable an epidemiologically-informed description of cognitive development that spans the newborn period through young adulthood. Together, the two studies will provide comprehensive information on the neural correlates of specific cognitive domains as they develop over time and novel insights into on how substance exposure at various developmental stages impacts discrete cognitive functions. As the HBCD study sample approaches later childhood, longitudinal findings from ABCD can be used to inform the specifics of future the HBCD study assessments and whether it would be fruitful for constructs such as risk-related decision-making to be examined at younger ages.

3.3. Item and test bias

The NCL Workgroup considered issues of bias in assessment decisions. For example, the decision was made to use the full MB-CDI rather than a shorter adaptive version as data were not available to determine if the adaptive algorithm performed equally across various participant dimensions (e.g., language spoken, income, race/ethnicity). The workgroup will monitor bias by including factors such as the language used to administer the task and the child's biological sex, ethnicity, and race in quality control analyses; which will inform any protocol adjustments. Several of the outcome measures have known differences associated with one or more of these factors (Lowe et al., 2019; Botdorf et al., 2022; Duncan et al., 2012) and it is not the intent of the NCL Workgroup activities to replicate or extend these findings. We welcome such use of the data to clarify exposomal factors that may contribute to differential outcomes and believe the large cohort collected across the United States will afford opportunities for valuable contributions in this area. Instead, the NCL Workgroup will focus on potential biases associated with the data collection procedures unique to the study to inform future users of the data.

The HBCD study is committed to enrolling families who speak English and Spanish. The NCL Workgroup recognizes the challenges and limitations of assessing bilingual language development. A direct translation of language items and verbal instructions from English to Spanish may not adequately capture the wide array of dialects and linguistic variations (Tellez et al., 2023) used by Spanish-speaking participants given the differences in translations across sites and the regional variations in Spanish dialects. Additionally, the structural differences between the Spanish and English language, such as the transformation of monosyllabic words in English into multi-syllabic in Spanish (e.g., "ball" versus "pelota") further complicates the matter. Using an English language-based testing tool could potentially introduce bias when assessing non-English speaking children, as highlighted in a prior study (Lowe et al., 2013) where children who had Spanish as the primary language within the home obtained significantly lower language scores compared to children with English as their primary language, despite similar neurocognitive scores between both groups. The developers of the NBT attempted to equate item difficulty in their language assessments, which may minimize this problem but it persists in the Bayley-4 and deferred imitation task assessments conducted in the study. The HBCD study offers a unique opportunity to explore assessment methods that are culturally appropriate for children from non-English speaking homes. The anticipated substantial sample size of the study presents potential opportunity for assessing item bias and differential predictive validity of infant assessment measures over time and has the potential to inform future bilingual assessment strategies.

4. Conclusions

The HBCD Study provides an unprecedented opportunity to examine early brain and behavioral development in a large cohort of US young children and their families. The open science framework and timely data releases will allow scientists across disciplines to contribute to knowledge of the factors affecting developmental outcomes. The assessments of neurocognition and language are of critical importance in linking the extensive neuroimaging (EEG and MRI) data to functional outcomes in this cohort. These data can be used to assess functional outcomes of independent variables collected in the study, to assess developmental changes and processes in the functional outcomes, and the interactions between these two. These outcomes will help with establishing the long-term effects of exposure to substances and other environmental, social and biological factors that may influence development.

The design of the HBCD study will guide how data can be used to answer scientific questions. The windows for visits cover a range of ages (i.e., 3–9 months, 9 – 15 months, etc.). During any given window, infants' ability to demonstrate their development varies greatly. The workgroup chose to maintain consistency of the constructs assessed during each visit, however the methods of assessment vary within a visit. Thus, while the task battery and/or methods or responding may vary within a visit – there is a standard battery based on the age of the infant. For example, infants who are assigned an early V03 window (< 6 months), will not complete gaze tasks as a result of differences in visual acuity and developmental changes in brain maturation related to attention and eye movement control. V06 visits also have variability in their response format of assessments across the age span. For example, the switch from eye gaze control of attention to using touchscreen methods to assess EF will result in potential response format variation both within a given visit window across participants and within individuals over time. This decision was made to allow maximal use of the data for trajectory modeling, while still enabling cross-sectional comparisons of age-normed scores (for example from the Mullen Scales).

The NCL Workgroup worked collaboratively with other workgroups within the consortium to integrate the assessment procedures to improve the overall efficiency of the HBCD study assessment battery. Extensive discussions were held with other workgroups to minimize the HBCD study staff and participant (caregiver and infant) burden. Reducing visit length was needed to enable the HBCD study staff to successfully maintain the proposed visit schedules and to prevent attrition for assessments at a given visit due to infant fatigue or over the course of the study due to parental distress. For example, gaze tasks used in the BTB were separated by activities that involved active participation and movement to minimize infant fatigue with watching the screen. Such tasks were drawn from other NCL Workgroup assessments and other workgroup assessments.

The selected NCL components of the HBCD study assessment battery reflect the NCL Workgroup's efforts to balance issues of participant burden, known health and socioeconomic disparities, psychometrics, and feasibility in compiling the assessment battery. To enable data collection of such a large cohort across diverse the HBCD study staffing models across sites, assessments relied on standardized scoring systems that reduce research assistant burden, limit the need for advanced assessment expertise, and enhanced data transfer. Early HBCD study data will be monitored to ensure high-quality data collection and will adapt as new discoveries inform the assessment methods.

The NCL Workgroup anticipates that the devised assessment protocol of cognitive and language development used in the HBCD study will enable meaningful contributions to early maturational brain changes related to the domains of function selected for study. Data collected in this study are anticipated to provide an improved understanding of exposomal and teratogenic factors that contribute to alterations of neurodevelopmental functions linked to maturational brain changes.

CRediT authorship contribution statement

Emily S Kuschner: Writing – review & editing, Writing – original draft. **Zoe Hulce:** Writing – review & editing, Writing – original draft, Visualization. **Eric H Zimak:** Writing – review & editing, Writing – original draft. **Sherry DeGray:** Writing – review & editing, Writing – original draft. **Jennifer Vannest:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization. **Lucia Ciciolla:** Writing – review & editing, Writing – original draft, Methodology. **Sidney Smith:** Writing – review & editing, Writing – original draft. **Stefanie C Bodison:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization. **So Yeon Shin:** Writing – review & editing, Methodology. **Patricia M Blasco:** Writing – review & editing. **Tracy Riggins:** Writing – review & editing, Writing – original draft, Validation, Data curation, Conceptualization. **Natacha Akshoomoff:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Data curation, Conceptualization. **Alexandra Perez:** Writing – review & editing, Writing – original draft. **Alexandra S Potter:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization. **Miriam A Novack:** Writing – review & editing, Writing – original draft, Resources. **Julie A Kable:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Monica Luciana:** Writing – review & editing, Writing – original draft, Methodology. **Britley Learnard:** Writing – review & editing, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

No data was used for the research described in the article.

Acknowledgements

Data/Processes/Plans/Concepts (select as appropriate) used in the preparation of this article were obtained from the Healthy Brain and Child Development (HBCD) Study (https://hbcdstudy.org/). This is a multisite, longitudinal study designed to recruit approximately 7500 families and follow them from pregnancy to early childhood. The HBCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA055352, U01DA055353, U01DA055366, U01DA055365, U01DA055362, U01DA055342, U01DA055360, U01DA055350, U01DA055338, U01DA055355, U01DA055363, U01DA055349, U01DA055361, U01DA055316, U01DA055344, U01DA055322, U01DA055369, U01DA055358, U01DA055371, U01DA055359, U01DA055354, U01DA055370, U01DA055347, U01DA055357, U01DA055367, U24DA055325, U24DA055330. A full list of supporters is available at https://hbcdstudy.org/about/federal-partners/. A listing of participating sites and a complete listing of the study investigators can be found at https://hbcdstudy.org/study-sites/. HBCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or HBCD consortium investigators. Additional support was provided to one of the authors from their work on the Adolescent Brain Cognitive Development Study (ABCD) (U01DA041120). A special thank you to Dr. Carol Cheatham, UNC, for design, creation and dissemination of the deferred imitation props and to Dr. Scott Johnson, UCLA, for his leadership in the first two years of the NCL Workgroup and his involvement in selecting the measures.

Dr. Michelle Freund was substantially involved in all of the cited grants in her capacity as Scientific Director for the HBCD Study, Drs. Katherine Cole and Katia Howlett were substantially involved in U01DA055369, consistent with their roles as Scientific Program Manager and Deputy Director of Division of Extramural Research and Co-Chairs of the Communications, Engagement and Dissemination Committee, respectively. All other Federal representatives contributed to the interpretation of the data and participated in the preparation, review and approval of the manuscript, consistent with their roles on the HBCD transNIH HBCD Collaborators. The views and opinions expressed in this manuscript are those of the authors only and do not necessarily represent the views, official policy or position of the U.S. Department of Health and Human Services or any of its affiliated institutions or agencies.

References

Allopenna, P.D., Magnuson, J.S., Tanenhaus, M.K., 1998. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. J. Mem. Lang. 38 (4), 419–439.

Amer, E.M., Abd El-Maksoud, G.M., Mahgoub, E.A., 2022. Assessment of sensory integrative function of elementary school children In Egypt. J. Occup. Ther., Sch., Early Interv. 15 (2), 218–230.

Amoros, R., et al., 2019. Maternal copper status and neuropsychological development in infants and preschool children. Int J. Hyg. Environ. Health 222 (3), 503–512.

Anokhin, A., et al., 2022. Age-related changes and longitudinal stability of individual differences in ABCD Neurocognition measures. *54, 101078*. Developmental Cognitive Neuroscience Dev. Cogn. Neurosci. 54, 101078.

Aylward, G.P., et al., 2022. Assessment of executive function in infants and toddlers: a potential role of the bayley-4. J. Dev. Behav. Pedia 43 (7), e431–e441.

Bakhireva, L.N., et al., 2019. Association between prenatal opioid exposure, neonatal opioid withdrawal syndrome, and neurodevelopmental and behavioral outcomes at 5-8 months of age. Early Hum. Dev. 128, 69–76.

Bauer, P.J., 2010. Varieties of early experience: implications for the development of declarative memory in infancy. Preface. Adv. Child Dev. Behav. 38 (p. xi-xiii).

Baum, S.H., Stevenson, R.A., Wallace, M.T., 2015. Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. Prog. Neurobiol. 134, 140–160.

Bayley, N., 2023. Escalas Bayley de desarrollo infantil | Tercera Edición, Spain Version. Pearson, Bloomington, MN.

Bayley, N., Aylward, G., 2019a. Administration Manual: Bayley Scales of Infant and Toddler Development, Fourth Edition. Pearson, Bloomington, MN.

Bayley, N., Aylward, G., 2019b. Bayley Scales of Infant and Toddler Development Fourth Edition: Technical Manual (Vol. Bloomington, MN). Pearson.

Beker, F., et al., 2023. Smell and taste of milk during tube feeding of preterm infants: neurodevelopmental follow-up of the randomized TASTE trial, study protocol. Trials 24 (1), 290.

Bell, M.A., 1998. Frontal lobe function during infancy: Implications for the development of cognition and attention. In: Richards, J. (Ed.), Cognitive Neuroscience of Attention: A Developmental Perspective. Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey, pp. 287–316.

Biagas, K.V., et al., 2023. Scoping Review: Neurocognitive Outcome Assessments After Critical Illness in Children. J. Intensive Care Med 38 (4), 358–367.

Bishop, S.L., et al., 2011. Convergent validity of the Mullen Scales of Early Learning and the differential ability scales in children with autism spectrum disorders. Am. J. Intellect. Dev. Disabil. 116 (5), 331–343.

Blazhenkova, O., Kozhevnikov, M., 2010. Visual-object ability: A new dimension of non-verbal intelligence. Cognition 117 (3), 276–301.

Botdorf, M., et al., 2022. Socioeconomic disadvantage and episodic memory ability in the ABCD sample: Contributions of hippocampal subregion and subfield volumes. *57, 101138. https://doi.org/10.1016/j.dcn.2022.101138.* Developmental cognitive neuroscience Dev. Cogn. Neurosci. 57, 101138.

Brand, M., et al., 2002. The game of dice - a new test for the assessment of risktaking behavior. Neurorehabilitation Neural Repair 16, 142–143.

Cattani, A., et al., 2010. Communicative and linguistic development in preterm children: a longitudinal study from 12 to 24 months. Int J. Lang. Commun. Disord. 45 (2), 162–173.

Charkaluk, M.L., et al., 2019. Association of language skills with other developmental domains in extremely, very, and moderately preterm children: EPIPAGE 2 cohort study. J. Pedia 208, 114–120 e5.

Chawla, N., et al., 2009. Perceived approval of friends and parents as mediators of the relationship between self-determination and drinking. J. Stud. Alcohol Drugs 70 (1), 92–100.

Coles, C.D., et al., 2019. Gestational age and socioeconomic status as mediators for the impact of prenatal alcohol exposure on development at 6 months. Birth Defects Res 111 (12), 789–796.

Dack, K., et al., 2022. Prenatal mercury exposure and neurodevelopment up to the age of 5 years: a systematic review. Int J. Environ. Res Public Health 19 (4).

Daughters, S.B., et al., 2009. Distress tolerance and early adolescent externalizing and internalizing symptoms: the moderating role of gender and ethnicity. Behav. Res Ther. 47 (3), 198–205.

Davis, B., 2015. Spatial reasoning in the early years: principles, assertions, and speculations. Routledge, New York viii, 176 pages.

DeAnda, S., et al., 2016. The language exposure assessment tool: quantifying language exposure in infants and children. J. Speech Lang. Hear Res 59 (6), 1346–1356.

Del Rosario, C., et al., 2021. How to use the Bayley Scales of Infant and Toddler Development. Arch. Dis. Child Educ. Pr. Ed. 106 (2), 108–112.

Despres, C., et al., 2005. Neuromotor functions in Inuit preschool children exposed to Pb, PCBs, and Hg. Neurotoxicol. Teratol. 27 (2), 245–257.

Diamond, A., 2020. Executive functions. Handb. Clin. Neurol. 173, 225–240.

Dikmen, S.S., et al., 2014. Measuring episodic memory across the lifespan: NIH Toolbox Picture Sequence Memory Test. J. Int Neuropsychol. Soc. 20 (6), 611–619.

Doherty, B.T., et al., 2019. Prenatal exposure to organophosphate esters and cognitive development in young children in the Pregnancy, Infection, and Nutrition Study. Environ. Res 169, 33–40.

Duncan, A.F., et al., 2012. Effect of ethnicity and race on cognitive and language testing at age 18-22 months in extremely preterm infants. J. Pedia 160 (6), 966–971 e2.

Dunn, W., 2007. Supporting children to participate successfully in everyday life by using sensory processing knowledge. Infants Young-.-. Child. 20 (2), 84–101.

Dunn, W., et al., 2022. Construct Validity of the Sensory Profile Interoception Scale: Measuring Sensory Processing in Everyday Life. Front Psychol. 13, 872619.

Fenson, L., et al., 1993. MacArthur Communicative Development Inventories: User's Guide and Technical Manual (Baltimore, MD: Brookes Publishing Co). Brookes Publishing Co, Baltimore, MD.

Fenson, L., Marchman, V.A., Thal, D.J., Dale, P.S., Reznick, J.S., and Bates, E. (2007). MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual, 2nd Edn. Baltimore, MD: Brookes Publishing Co. doi: 10.1037/t11538-000 MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual, 2nd Edn. 2007, Baltimore, MD: Brookes Publishing Co.

Fernández-Pires, P., et al., 2021. Sleep duration and quality and sensory reactivity in school-aged children: the spanish cross-sectional InProS Study. Front. Pediatr. 9.

Fjeldsted, B., Xue, L., 2019. Sensory processing in young children with fetal alcohol spectrum disorder. Phys. Occup. Ther. Pedia 39 (5), 553–565.

Foster-Cohen, S., et al., 2007. Early delayed language development in very preterm infants: evidence from the MacArthur-Bates CDI. J. Child Lang. 34 (3), 655–675.

Fryer, S.L., et al., 2007. Prenatal alcohol exposure affects frontal-striatal BOLD response during inhibitory control. Alcohol Clin. Exp. Res 31 (8), 1415–1424.

Gershon, R.C., et al., 2013. IV. NIH Toolbox Cognition Battery (CB): measuring language (vocabulary comprehension and reading decoding). Monogr. Soc. Res Child Dev. 78 (4), 49–69.

Gershon, R.C., et al., 2014. Language measures of the NIH Toolbox Cognition Battery. J. Int Neuropsychol. Soc. 20 (6), 642–651.

Gershon, R., M. Novack., and A. Kaat, *The NIH infant and Toddler Toolbox: A new standardized tool for assessing neurodevelopment in children ages 1-42 months.* 2024.

Guillaume, M., et al., 2023. Groupitizing reflects conceptual developments in math cognition and inequities in math achievement from childhood through adolescence. Child Dev. 94 (2), 335–347.

Hendricks, G., et al., 2020. Prenatal alcohol exposure is associated with early motor, but not language development in a South African cohort. Acta Neuropsychiatr. 32 (3), 1–8.

Hernandez-Martinez, C., et al., 2017. Effects of prenatal nicotine exposure on infant language development: a cohort follow up study. Matern Child Health J. 21 (4), 734–744.

Jackson-Maldonado, D., et al., 2003. MacArthur inventarios del desarrollo de habilidades communicativas: user's guide and technical manual (Baltimore, MD: Brookes Publishing Co). Brookes Publishing Co, Baltimore, MD.

Jirikowic, T.L., et al., 2013. Sensory control of balance: a comparison of children with fetal alcohol spectrum disorders to children with typical development. J. Popul Ther. Clin. Pharm. 20 (3), e212–e228.

Johnson, S.P., Moore, D.S., 2020. Spatial thinking in infancy: origins and development of mental rotation between 3 and 10 months of age. Cogn. Res Princ. Implic. 5 (1), 10.

Kauley, N., et al., 2024. Predicting communication skills outcomes for preschool children with autism spectrum disorder following early intervention. Neuropsychiatr. Dis. Treat. 20, 35–48.

Kern, J.K., et al., 2015. Are ASD and ADHD a Continuum? A Comparison of Pathophysiological Similarities Between the Disorders. J. Atten. Disord. 19 (9), 805–827.

Kfir, M., et al., 2009. Can prenatal ultrasound detect the effects of in-utero alcohol exposure? A pilot study. Ultrasound Obstet. Gynecol. 33 (6), 683–689.

Khoury, J.E., Milligan, K., Girard, T.A., 2015. Executive functioning in children and adolescents prenatally exposed to alcohol: a meta-analytic review. Neuropsychol. Rev. 25 (2), 149–170.

Lauer, J.E., Lourenco, S.F., 2016. Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. Psychol. Sci. 27 (10), 1291–1298.

Lees, B., et al., 2020. Association of prenatal alcohol exposure with psychological, behavioral, and neurodevelopmental outcomes in children from the adolescent brain cognitive development study. Am. J. Psychiatry 177 (11), 1060–1072.

Lowe, J.R., et al., 2013. Effect of primary language on developmental testing in children born extremely preterm. Acta Paediatr. 102 (9), 896–900.

Lowe, J.R., et al., 2019. Behavioral problems are associated with cognitive and language scores in toddlers born extremely preterm. Early Hum. Dev. 128, 48–54.

Lucas, B.R., et al., 2014. Gross motor deficits in children prenatally exposed to alcohol: a meta-analysis. Pediatrics 134 (1), e192–e209.

Luciana, M., et al., 2018. Adolescent neurocognitive development and impacts of substance use: overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery. Dev. Cogn. Neurosci. 32, 67–79.

Lund, J.I., et al., 2020. Adverse childhood experiences and executive function difficulties in children: a systematic review. Child Abus. Negl. 106, 104485.

Mahabee-Gittens, E.M., et al., 2024. Prenatal opioid exposure and risk for adverse brain and motor outcomes in infants born premature. J. Pedia 267, 113908.

Marban-Castro, E., et al., 2022. Neurodevelopment in normocephalic children exposed to zika virus in utero with no observable defects at birth: a systematic review with meta-analysis. Int J. Environ. Res Public Health 19 (12).

Marchman, V.A., Dale, P.S., 2023. The macarthur-bates communicative development inventories: updates from the CDI advisory board. Front Psychol. 14, 1170303.

Marchman, V.A., Dale, P.S., Fenson, L., Dale, P.S., Fenson, L., 2023. In: Marchman, V.A. V.A., Dale, P.S., Fenson, L. (Eds.), The MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual, 3rd edition. Brookes Publishing Co, Baltimore.

Mullen, E.M., Infant MSEL manual: Infant Mullen Scales of Early Learning. 1991, Cranston, R.I. (244 Deerfield Road, Cranston 02920): T.O.T.A.L. Child, Inc. ix, 115 p.

Narzisi, A., et al., 2022. Sensory profiles in school-aged children with autism spectrum disorder: a descriptive study using the sensory processing measure-2 (SPM-2). J. Clin. Med 11 (6).

Nelson, C.A., 2007. A Neurobiological Perspective on Early Human Deprivation. Child Dev. Perspect. 1 (1), 13–18.

Nelson, C.A., Frankeberger, J., Chambers, C.D., 2024. An introduction to the HEALthy brain and child development (HBCD) study. Dev. Cogn. Neurosci., 101441

NIH Infant and Toddler (Baby) Toolbox. June 24,2022 [cited 2023; Available from: 〈https://neuroscienceblueprint.nih.gov/resources-tools/blueprint-resources-tools-library/nih-infant-and-toddler-baby-toolbox〉.

O'Hare, E.D., et al., 2009. Altered frontal-parietal functioning during verbal working memory in children and adolescents with heavy prenatal alcohol exposure. Hum. Brain Mapp. 30 (10), 3200–3208.

Olegard, R., et al., 1979. Effects on the child of alcohol abuse during pregnancy. Retrospective and prospective studies. Acta Paediatr. Scand. Suppl. 275, 112–121.

Olesen, T.S., et al., 2018. Prenatal phthalate exposure and language development in toddlers from the Odense Child Cohort. Neurotoxicol Teratol. 65, 34–41.

Owens, M.M., et al., 2022. One-year predictions of delayed reward discounting in the adolescent brain cognitive development study. Exp. Clin. Psychopharmacol. 30 (6), 928–946.

Oyungu, E., et al., 2022. Predicting neurodevelopmental risk in children born to mothers living with HIV in Kenya: protocol for a prospective cohort study (Tabiri Study). BMJ Open 12 (4), e061051.

Parham, L.D., et al., 2021. Sensory processing measure, second edition. Western Psychological Services. *(SPM-2)*.

Patil, O., Kaple, M., 2023. Sensory processing differences in individuals with autism spectrum disorder: a narrative review of underlying mechanisms and sensory-based interventions. Cureus 15 (10), e48020.

Piaget, Jean (ed). Play, Dreams and Imitation in Childhood. ed. G. S and H. FM. 1962, Routledge & Kegan Paul: London.

Richmond, J., Colombo, M., Hayne, H., 2007. Interpreting visual preferences in the visual paired-comparison task. J. Exp. Psychol.: Learn., Mem., Cogn. 33 (5), 823–831.

Riggins, T., et al., 2013. Elicited imitation performance at 20 months predicts memory abilities in school-age children. J. Cogn. Dev. 14 (4), 593–606.

Rittle-Johnson, B., Zippert, E.L., Boice, K.L., 2019. The roles of patterning and spatial skills in early mathematics development. Early Child. Res. Q. 46, 166–178.

Román-Oyola, R., Reynolds, S., 2013. Prevalence of sensory modulation disorder among puerto rican preschoolers: an analysis focused on socioeconomic status variables. Occup. Ther. Int. 20 (3), 144–154.

Rosenberg, S.J., Ryan, J.J., Prifitera, A., 1984. Rey auditory-verbal learning test performance of patients with and without memory impairment. J. Clin. Psychol. 40 (3), 785–787.

Ruff, H.A., Rothbart, M.K., 1996. *Attention in early development: Themes and variations.* Attention in early development: Themes and variations. . Oxford University Press, New York, NY US.

Smolker, H.R., et al., 2022. The Emotional Word-Emotional Face Stroop task in the ABCD study: Psychometric validation and associations with measures of cognition and psychopathology. Dev. Cogn. Neurosci. 53, 101054.

Sowell, E.R., et al., 2007. Functional magnetic resonance imaging of verbal learning in children with heavy prenatal alcohol exposure. Neuroreport 18 (7), 635–639.
Sparrow, S.S., Cicchetti, D.V., Balla, D.A., 2005. Vineland Adaptive Behavior Scales. Survey Forms Manual, 2nd edition. American Guidance Service, Circle Pines, Minn.
Sparrow, S.S., Cicchetti, D.V., & Saulnier, C.A. (2016). Vineland Adaptive Behavior Scales, Third Edition (Vineland-3). San Antonio, TX: Pearson., Vineland Adaptive Behavior Scales, Third Edition (Vineland-3). 2016, San Antonio, TX: Pearson.
Subramoney, S., et al., 2018. The early developmental outcomes of prenatal alcohol exposure: a review. Front Neurol. 9, 1108.
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. Psychol. Assess. 27 (4), 1364–1378.
Tellez, E.S., et al., 2023. Regionalized models for Spanish language variations based on Twitter. Lang. Resour. Eval. 57 (4), 1697–1727.
Thompson, W.K., et al., 2019. The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: findings from the ABCD study's baseline neurocognitive battery. Dev. Cogn. Neurosci. 36, 100606.
Volkow, N.D., et al., 2018. The conception of the ABCD study: From substance use to a broad NIH collaboration. Dev. Cogn. Neurosci. 32, 4–7.
Warren, K., et al. *Consensus Statement on FASD*. FASD Terminology Summit 2004 03/01/2011]; Available from: ⟨http://www.nofas.org/advocate/terminology.aspx⟩.
Wechsler D. (2003). Wechsler Intelligence Scale for Children, t.E.S.A., TX: PsychCorp., *Wechsler Intelligence Scale for Children,* 4th Ed. 2003, San Antonia, TX: PsychCorp.
Weintraub, S., et al., 2013a. I. NIH Toolbox Cognition Battery (CB): introduction and pediatric data. Monogr. Soc. Res Child Dev. 78 (4), 1–15.
Weintraub, S., et al., 2013b. Cognition assessment using the NIH Toolbox. Neurology 80 (11 Suppl 3) p. S54-64.
Wozniak, J.R., et al., 2013. Global functional connectivity abnormalities in children with fetal alcohol spectrum disorders. Alcohol.: Clin. Exp. Res. 37 (5), 748–756.
Yeoh, S.L., et al., 2019. Cognitive and motor outcomes of children with prenatal opioid exposure: a systematic review and meta-analysis. JAMA Netw. Open 2 (7), e197025.
Zimmerman, I.L., Steiner, V.G., Pond, R.E., 2002. Preschool Language Scale, 4th ed.). The Psychological Corporation, San Antonio, TX.